Form PTO 1390 (REV 5-93) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER B45168
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED / ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5) 09/868604
INTERNATIONAL APPLICATION NO PCT/EP99/10297	INTERNATIONAL FILING DATE 21 December 1999	PRIORITY DATE CLAIMED 21 December 1998
TITLE OF INVENTION VACCINE		
APPLICANT(S) FOR DO/EO/US Alex BOLLEN, Alain FAUCONNIER, and Edmond GODFROID		

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1 [x] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.

2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.

3. [x] This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).

4. [x] A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.

5. [x] A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. [ ] is transmitted herewith (required only if not transmitted by the International Bureau).
   b. [x] has been transmitted by the International Bureau.
   c. [ ] is not required, as the application was filed in the United States Receiving Office (RO/US).

6. [ ] A translation of the International Application into English (35 U.S.C. 371(c)(2)).

7. [x] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. [ ] are transmitted herewith (required only if not transmitted by the International Bureau).
   b. [x] have been transmitted by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.

8. [ ] A translation of the amendments to the claims under PCT Article 19 (35 U.S. C. 371(c)(3)).

9. [ ] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).

10. [ ] A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern other document(s) or information included:

11. [x] An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98; and Form PTO-1449.

12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. 3.28 and 3.31 is included.

13. [x] A FIRST preliminary amendment.

14. [ ] A SECOND or SUBSEQUENT preliminary amendment.

15. [x] Please amend the specification by inserting before the first line the sentence: This is a 371 of International Application PCT/EP99/10297, filed December 21, 1999, which claims benefit from the following Provisional Application: GB 9828217.1 filed 21 December 1998.

16. [ ] A substitute specification.

17. [ ] A change of power of attorney and/or address letter.

18. [x] An Abstract on a separate sheet of paper.

19. [x] Other items or information: Sequence Listing, Statement to Support, Diskette

JC18 Rec'd PCT/PTO 2 0 JUN 2001

US APPLICATION NO. (if known see 37 CFR 1.5) 09/868604	INTERNATIONAL APPLICATION NO. PCT/EP99/10297	ATTORNEYS DOCKET NO B45168

20. [X] The following fees are submitted:				CALCULATIONS	PTO USE ONLY
Basic National Fee (37 C.F.R. 1.492(a)(1)-(5)):					
Search Report has been prepared by the EPO or JPO **$860.00**					
International Preliminary Examination Fee paid to USPTO (37 CFR 1.482) **$690.00**					
No International Preliminary Examination Fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) **$710.00**				**$710.00**	
Neither International Preliminary Examination Fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....**$1,000.00**					
International Preliminary Examination Fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4).....**$100.00**					
ENTER APPROPRIATE BASIC FEE AMOUNT =				**$710.00**	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				**$0.00**	
Claims	Number Filed	Number Extra	Rate		
Total claims	**99 - 20 =**	**79**	**79 x $18.00**	**$1422.00**	
Independent claims	**10 - 3 =**	**7**	**7 x $80.00**	**$560.00**	
Multiple dependent claims (if applicable)			**+ $270.00**	**$270.00**	
TOTAL OF ABOVE CALCULATIONS =				**$2252.00**	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).				**$**	
SUBTOTAL =				**$2962.00**	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)) +				**$**	
TOTAL NATIONAL FEE =				**$2962.00**	
				Amount to be refunded	$
				charged	$

a. ☐ A check in the amount of $____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 19-2570 in the amount of **$2962.00** to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-2570 . A duplicate copy of this sheet is enclosed.

d. ☒ General Authorization to charge any and all fees under 37 CFR 1.16 or 1.17, including petitions for extension of time relating to this application (37 CFR 1.136 (a)(3)).

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
GLAXOSMITHKLINE
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, PA 19406-0939
Phone (610) 270-5024
Facsimile (610) 270-5090

SIGNATURE
Zoltan Kerekes
NAME
38,938
REGISTRATION NO.

n:\zk\apps\b45168\us natl.doc

Attorney Docket No. B45168

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Bollen, et al.	20 June 2001
Int'l. App. No.:	PCT/EP99/10297	Group Art Unit: Not Yet Assigned
Int'l. Filing Date:	21 December 1999	Examiner: Not Yet Assigned
For:	VACCINE	

Assistant Commissioner of Patents
Box: PCT
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Preliminary to the examination of this application, Applicants respectfully request consideration and entry of the following Preliminary Amendment.

Applicants are submitting herewith a new Statement to Support Filing and Submission in Accordance with 37 CFR §§ 1.821 Through 1.825, which includes three corrected sheets (pages 111, 113, and 114) pursuant to 37 CFR §§ 1.825. In addition, Applicants are submitting the complete Sequence Listing on computer diskette.

IN THE SPECIFICATION

Please amend Table 3 on page 38 as follows:

Table 3

Names	Coding sequence from/to (with reference to Fig. 5)	Coding DNA strand	SEQ ID NO:	Homologous genes (from *Yersinia*, unless otherwise specified)
Class II ORFs which putatively code for effector proteins				
bopN	11906/13003	complement	41	*YopN* (= *lcrE*)
orf1	6160/6747	direct	43	None
orf2	10752/11120	complement	45	None

orf3	11117/11527	complement	47	None
orf4	11532/11909	complement	49	None
orf5	13002/13784	direct	51	None
orf6	13806/14081	direct	53	None
orf7	14630/15571	direct	55	None
orf8	15601/16803	direct	57	None
orf9	16827/17288	direct	59	*BcrH*
orf10	17293/17814	direct	61	*pcr4* (*Pseudomonas aeruginosa*)
orf11	29412/29591	complement	63	None
orf12	29555/30529	complement	65	None
orf13	30631/31776	direct	67	None
orf14	31818/33005	complement	69	None
orf15	32370/33014	direct	71	None

IN THE CLAIMS:

Please cancel claims 1-29.

Please add new claims 30-78.

30. An isolated polypeptide comprising an amino acid sequence which has at least 75% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 and 72 over its entire length.

31. The polypeptide as claimed in claim 30 comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 and 72.

32. An isolated polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72.

33. An isolated polypeptide comprising a fragment of at least 7 consecutive amino acids of the polypeptide as claimed in any one of claims 30 to 32, wherein the fragment comprises an epitope.

34. The polypeptide of claim 33, wherein the fragment is immunogenic.

35. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 75% identity to the amino acid sequence of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72 over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

36. An isolated polynucleotide comprising a nucleotide sequence that has at least 75% identity to a nucleotide sequence, encoding a polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72, over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

37. An isolated polynucleotide which comprises a nucleotide sequence which has at least 75% identity to that of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

38. The isolated polynucleotide as claimed in claim 35 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

39. The isolated polynucleotide as claimed in claim 36 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

40. The isolated polynucleotide as claimed in claim 37 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

41. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72.

42. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71.

43. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72, obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 or a fragment thereof.

44. An expression vector comprising an isolated polynucleotide according to any one of claims 35-43.

45. A recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 35-43.

46. A host cell comprising the expression vector of claim 44 or a subcellular fraction or a membrane of said host cell.

47. A process for producing the polypeptide of claim 30 comprising the steps of culturing a host cell of claim 46 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

48. A process for expressing a polynucleotide of any one of claims 35-43 comprising transforming a host cell with an expression vector comprising at least one of said polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

49. A vaccine composition comprising an effective amount of the polypeptide of claim 30 and a pharmaceutically acceptable carrier.

50. A vaccine composition comprising an effective amount of the polypeptide of claim 31 and a pharmaceutically acceptable carrier.

51. A vaccine composition comprising an effective amount of the polypeptide of claim 32 and a pharmaceutically acceptable carrier.

52. A vaccine composition comprising an effective amount of the polypeptide of claim 33 and a pharmaceutically acceptable carrier.

53. A vaccine composition comprising an effective amount of the polypeptide of claim 34 and a pharmaceutically acceptable carrier.

54. The vaccine composition of claim 49, wherein the polypeptide has an amino acid sequence selected from the group consisting of: SEQ ID NO:42, 46, 48, 50, 52, 54, 56, 58, 60 and 62.

55. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 35 to 43 and a pharmaceutically acceptable carrier.

56. The vaccine composition according to any one of claims 49-55, wherein said composition comprises at least one other *Bordetella pertussis* antigen.

57. An antibody immunospecific for the amino acid sequence of claim 30 or 31.

58. An antibody immunospecific for the polypeptide of claim 32.

59. An antibody immunospecific for the fragment of claim 33.

60. An antibody immunospecific for the fragment of claim 34.

61. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 30, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

62. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 31, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

63. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 32, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

64. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 33, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

65. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 34, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

66. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 30 and a suitable pharmaceutical carrier.

67. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 31 and a suitable pharmaceutical carrier.

68. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 32 and a suitable pharmaceutical carrier.

69. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 33 and a suitable pharmaceutical carrier.

70. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 34 and a suitable pharmaceutical carrier.

71. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polynucleotide of claims 35-43.

72. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 30.

73. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 31.

74. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 32.

75. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 33.

76. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 34.

77. A method of identifying virulence genes from a pathogenicity island containing a type III secretion system from pathogenic strains of bacteria, comprising:

designing degenerate PCR primers complementary to well-conserved regions specific to the LcrD polypeptide of *Yersinia*;

amplifying the polynucleotide containing the DNA sequence between (and including the DNA sequence of) the primers of *lcrD*-like genes present in said pathogenic strain of bacteria;

sequencing the *lcrD*-like gene;

determining whether the DNA sequence is more homologous: to the virulence-associated family of *lcrD*-like genes, or to the flagellar-associated family of *lcrD*-like genes; and

if a virulence-associated member, sequencing the entire pathogenicity island, and identifying genes within this sequence.

78. A method of determining whether a particular bacterial strain harbours a type III secretion system involved in pathogenicity, comprising:

designing degenerate PCR primers complementary to well-conserved regions specific to the LcrD polypeptide of *Yersinia*;

amplifying the polynucleotide containing the DNA sequence between (and including the DNA sequence of) the primers to determine the presence of any *lcrD*-like genes in said bacterial strain;

if amplified successfully, sequencing the *lcrD*-like gene; and

determining whether the DNA sequence is more homologous: to the virulence-associated family of *lcrD*-like genes, or to the flagellar-associated family of *lcrD*-like genes.

REMARKS

The above-identified application is being entered into the National Phase from PCT application no. PCT/EP99/10297.

Specification and Sequence Listing

An inadvertant error on page 38 (Table 3) which recites position 31773 as the end of the open reading frame for *orf14* has been corrected to show position 31818. As set forth below, correcting this error is obvious.

Applicants respectfully request amendment of SEQ ID NO:69 and SEQ ID NO:70 as set forth on substitute sheets 111, 113, and 114 submitted herewith pursuant to 37 CFR §§ 1.825.

In the original submission of the Sequence Listings, the Sequence Listing program completely ignored the stop codon at position 396. Although the sequence shown in Fig. 5 is correct, Table 3 states that *orf14* is encoded in the complementary strand from position 31773-33005. Although position 33005 indicates the correct start codon for this open reading frame, the end of the open reading frame is incorrectly stated. It should actually be at position 31818 – where the first stop codon (at position 396) is encountered.

Correcting this error is obvious. Given a properly stated start codon, and a correct DNA sequence, anyone would realize that the open reading frame MUST end where the first, in-frame,

stop codon in the sequence is encountered. This is obviously at position 396, and was clearly the intention of the Applicants at the time of filing the international application.

Three replacement sheets are provided as required under 37 CFR §§ 1.825. The error on page 111 has been rectified to indicate the number of nucleotides in SEQ ID NO:69 as being 1188. The error on page 113 has been rectified to indicate that the open reading frame does not extend past the stop codon at position 396. In addition, the number of amino acids in SEQ ID NO:70 has been rectified as being 395. The error on page 114 has been rectified to indicate the last amino acid in the open reading frame as being His395.

The corrected sequence listing for SEQ ID NO:69 and 70 does not go beyond the disclosure apparent to anyone from Fig. 5 and Table 3 of the International Application as filed.

The complete sequence listing is provided on a computer diskette. It is identical to the original written sequence listing in conjunction with the aforementioned corrections to SEQ ID NO: 69 and 70.

Claims

Claims 1-29 were cancelled. New claims 30-78 were added for the following reason: to put the claims in conformity with U.S. practice.

No new matter has been introduced.

Attached hereto is a marked-up version of the changes made to the specification and claims by the current amendment. The attached page is captioned **"Version with Markings to Show Changes Made"**. Applicants respectfully request that a timely Notice of Allowance be issued in this case.

Respectfully submitted,

Zoltan Kerekes
Attorney for Applicants
Registration No. 38,938

GLAXOSMITHKLINE
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, PA 19406-0939
Phone (610) 270-5024
Facsimile (610) 270-5090
N:\zk\apps\b45168\Preliminary Amendment.doc

VERSION WITH MARKINGS TO SHOW CHANGES

IN THE SPECIFICATION:

Table 3 appearing on page 38 has been amended as follows:

Table 3

Names	Coding sequence from/to (with reference to Fig. 5)	Coding DNA strand	SEQ ID NO:	Homologous genes (from *Yersinia*, unless otherwise specified)
Class II ORFs which putatively code for effector proteins				
bopN	11906/13003	complement	41	*YopN* (= *lcrE*)
orf1	6160/6747	direct	43	None
orf2	10752/11120	complement	45	None
orf3	11117/11527	complement	47	None
orf4	11532/11909	complement	49	None
orf5	13002/13784	direct	51	None
orf6	13806/14081	direct	53	None
orf7	14630/15571	direct	55	None
orf8	15601/16803	direct	57	None
orf9	16827/17288	direct	59	*BcrH*
orf10	17293/17814	direct	61	*pcr4* (*Pseudomonas aeruginosa*)
orf11	29412/29591	complement	63	None
orf12	29555/30529	complement	65	None
orf13	30631/31776	direct	67	None
orf14	[31773]31818/33005	complement	69	None
orf15	32370/33014	direct	71	None

IN THE SEQUENCE LISTING:

IN THE CLAIMS:

Claims 1-29 have been cancelled. New claims 30-78 have been added as follows:

30. An isolated polypeptide comprising an amino acid sequence which has at least 75% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 and 72 over its entire length.

31. The polypeptide as claimed in claim 30 comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 and 72.

32. An isolated polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72.

33. An isolated polypeptide comprising a fragment of at least 7 consecutive amino acids of the polypeptide as claimed in any one of claims 30 to 32, wherein the fragment comprises an epitope.

34. The polypeptide of claim 33, wherein the fragment is immunogenic.

35. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 75% identity to the amino acid sequence of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72 over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

36. An isolated polynucleotide comprising a nucleotide sequence that has at least 75% identity to a nucleotide sequence, encoding a polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72, over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

37. An isolated polynucleotide which comprises a nucleotide sequence which has at least 75% identity to that of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

38. The isolated polynucleotide as claimed in claim 35 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

39. The isolated polynucleotide as claimed in claim 36 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

40. The isolated polynucleotide as claimed in claim 37 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

41. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72.

42. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71.

43. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72, obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 or a fragment thereof.

44. An expression vector comprising an isolated polynucleotide according to any one of claims 35-43.

45. A recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 35-43.

46. A host cell comprising the expression vector of claim 44 or a subcellular fraction or a membrane of said host cell.

47. A process for producing the polypeptide of claim 30 comprising the steps of culturing a host cell of claim 46 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

48. A process for expressing a polynucleotide of any one of claims 35-43 comprising transforming a host cell with an expression vector comprising at least one of said polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

49. A vaccine composition comprising an effective amount of the polypeptide of claim 30 and a pharmaceutically acceptable carrier.

50. A vaccine composition comprising an effective amount of the polypeptide of claim 31 and a pharmaceutically acceptable carrier.

51. A vaccine composition comprising an effective amount of the polypeptide of claim 32 and a pharmaceutically acceptable carrier.

52. A vaccine composition comprising an effective amount of the polypeptide of claim 33 and a pharmaceutically acceptable carrier.

53. A vaccine composition comprising an effective amount of the polypeptide of claim 34 and a pharmaceutically acceptable carrier.

54. The vaccine composition of claim 49, wherein the polypeptide has an amino acid sequence selected from the group consisting of: SEQ ID NO:42, 46, 48, 50, 52, 54, 56, 58, 60 and 62.

55. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 35 to 43 and a pharmaceutically acceptable carrier.

56. The vaccine composition according to any one of claims 49-55, wherein said composition comprises at least one other *Bordetella pertussis* antigen.

57. An antibody immunospecific for the amino acid sequence of claim 30 or 31.

58. An antibody immunospecific for the polypeptide of claim 32.

59. An antibody immunospecific for the fragment of claim 33.

60. An antibody immunospecific for the fragment of claim 34.

61. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 30, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

62. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 31, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

63. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 32, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

64. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 33, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

65. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed claim 34, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

66. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 30 and a suitable pharmaceutical carrier.

67. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 31 and a suitable pharmaceutical carrier.

68. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 32 and a suitable pharmaceutical carrier.

69. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 33 and a suitable pharmaceutical carrier.

70. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claim 34 and a suitable pharmaceutical carrier.

71. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polynucleotide of claims 35-43.

72. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 30.

73. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 31.

74. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 32.

75. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 33.

76. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polypeptide of claim 34.

77. A method of identifying virulence genes from a pathogenicity island containing a type III secretion system from pathogenic strains of bacteria, comprising:

designing degenerate PCR primers complementary to well-conserved regions specific to the LcrD polypeptide of *Yersinia*;

amplifying the polynucleotide containing the DNA sequence between (and including the DNA sequence of) the primers of *lcrD*-like genes present in said pathogenic strain of bacteria;

sequencing the *lcrD*-like gene;

determining whether the DNA sequence is more homologous: to the virulence-associated family of *lcrD*-like genes, or to the flagellar-associated family of *lcrD*-like genes; and

if a virulence-associated member, sequencing the entire pathogenicity island, and identifying genes within this sequence.

78. A method of determining whether a particular bacterial strain harbours a type III secretion system involved in pathogenicity, comprising:

designing degenerate PCR primers complementary to well-conserved regions specific to the LcrD polypeptide of *Yersinia*;

amplifying the polynucleotide containing the DNA sequence between (and including the DNA sequence of) the primers to determine the presence of any *lcrD*-like genes in said bacterial strain;

if amplified successfully, sequencing the *lcrD*-like gene; and

determining whether the DNA sequence is more homologous: to the virulence-associated family of *lcrD*-like genes, or to the flagellar-associated family of *lcrD*-like genes.

Abstract

This invention relates to a general method for detecting pathogenic strains of bacteria which harbour a type III secretion system. More particularly, this invention relates to the methods as applied to the pathogen *Bordetella pertussis*. Furthermore, the invention relates to newly identified polynucleotides within these regions, virulent polypeptides encoded by them and to the use of such polynucleotides and polypeptides, and to their production. More particularly the polynucleotides and polypeptides of the present invention relate to the virulent effector proteins associated with the type III secretion system of *Bordetella pertussis*, which are particularly suitable for vaccine purposes.

VACCINE

FIELD OF INVENTION

This invention relates to a general method for detecting pathogenic strains of bacteria that harbour a type III secretion system, and characterising regions of the chromosome of said strain where virulence genes reside. More particularly, this invention relates to the method as applied to the pathogen *Bordetella pertussis*. Furthermore, the invention relates to newly identified polynucleotides within these regions, virulent polypeptides encoded by them and to the use of such polynucleotides and polypeptides, and to their production.

BACKGROUND OF THE INVENTION

Type III secretion systems:

Pathogenic bacteria invade many different niches in a broad host range and cause a wide variety of syndromes. It is due to this fact that it was believed previously that each disease might be induced by a distinct molecular mechanism. However, the spectrum of such mechanisms is not as broad as first imagined; rather, bacteria exploit a number of common molecular tools to achieve a range of goals. Among these tools are type III secretion systems, which provide a means for bacteria to target virulence factors directly at host cells. These factors then tamper with host cell functions to the pathogens' benefit.

The type III export system is responsible for secretion of *Salmonella* and *Shigella* invasion and virulence factors, Enteropathogenic *Escherischia coli* (EPEC) signal transduction molecules, virulence factors in several plant pathogens (for instance *Xanthomonas campestris pv. vesicatoria* [Fenselau *et al.*, 1992]) and Yops proteins in *Yersinia*. Yops export mechanism has been the most intensively investigated type III secretion apparatus (see for instance: Allaoui *et al.*, 1994; Bergman *et al.*, 1994). In this system, more than 20 different Ysc/Lcr proteins, all encoded by the virulence plasmid pYV, are presumed to compose a secretion channel spanning the *Yersinia* cell envelope.

Besides these elements involved in the secretion machinary, the pYV plasmid codes for the Yops proteins which are the secreted substrates and appear as the actual effectors of virulence.

Comparative studies of type III secretion systems originating from different species reveal that the components of the secretion machinery are conserved (Gygi *et al.*, 1995; Bogdanove *et al.*, 1996). In addition, homologs have been found in determinants which take part in flagellar assembly, indicating that this secretion pathway may be involved in surface organelle biosynthesis (Ramakrishnan *et al.*, 1991).

In contrast, however, the secreted substrates share no similarities, except in few cases. Therefore, the abandoned concept of a distinct molecular mechanism corresponding to each disease could reappear at the level of effector proteins.

Pathogenicity island

Pathogenicity islands have emerged as a novel theme in the field of bacterial virulence. Although they can comprise type III secretion systems they do not exlusively do so.

Early in the search for virulence genes, it was observed that many of these genes resided on plasmids. However, numerous virulence genes were also found on the chromosome. Surprisingly, the chromosomal virulence genes are also often clustered in functionally related groups. Such groups of virulence genes gave rise to the concept of pathogenicity islands (Pais) which can be defined as compact, distinct genetic units carrying virulence genes. These units, often flanked by direct repeats, occupy large chromosomal regions (often > 30 kb) and are present in pathogenic strains, whilst being absent or sporadically distributed in less-pathogenic (or non-pathogenic) strains of a bacterial species. These DNA segments are frequently associated with tRNA genes

and/or insertion sequence (IS) elements at their boundaries. In addition, their G+C content often differs from that of host bacterial DNA, suggesting a foreign origin.

Pathogenicity islands have been discovered in an increasing number of bacterial pathogens, including different categories of *E. coli, Salmonella typhimurium, Yersinia* spp, *Helicobacter pylori*, *Vibrio cholera* etc.

The first intensively studied pathogenicity islands were Pai I and Pai II, which encode the haemolysin determinants of uropathogenic *E. coli*. These two Pais, are flanked by direct repeats and can be deleted from the chromosome at frequencies of 10^{-4}, resulting in non-virulent mutant strains. Another pathogenicity island of 35 kb has recently been identified on the chromosome of enteropathogenic *E. coli* (EPEC) and was found to encode all known determinants involved in the so-called "attaching and effacing" (AE) lesion formation. This region was therefore referred to as "locus of enterocyte effacing" (LEE). Despite the fact that uropathogenic and enteropathogenic *E. coli* cause completely different infectious diseases, Pai I of the uropathogenic strains and the LEE locus of EPEC are inserted at exactly the same positions into the *E. coli* chromosome.

While some authors support a definition of pathogenicity islands which necessarily includes its chromosomal location, others have extended the concept to blocks of virulence genes, regardless of their location in chromosomes, plasmids or phages. The fact that, on one hand, phages and plasmids can easily insert into and excise from the chromosome and, on the other, that cryptic origins of plasmid replication, or phage related sequences were detected in Pais, prompted the latter and less restrictive definition.

The pathogenicity islands (PAIs) which code for a type III secretion system encompass genes that divide into two classes, I and II. Class I encompasses the genes coding for the secretion machinery components and their regulators of expression, class

II encompasses the genes encoding secreted effector proteins. Both *Yersinia lcrD* and *yscU* belong to class I. The precise functions of class I determinants is not well understood. Although it is sometimes not straightforward to make a clear distinction between class I and class II components, genes of class I can be identified as being present in many different species, and a comparison of their respective gene sequences indicate that equivalent genes share a significant (*yscI, yscO*) or even high level (*lcrD, yscU*, y*scN*) of sequence similarity (Hueck, 1998).

The second class of genes (class II) codes for proteins which constitute the substrate secreted by the translocon. These proteins appear as the actual effectors of virulence and are referred to as target proteins, virulence effector proteins or, simply, effectors. In contrast to the situation prevailing in class I gene products, the effectors share no, or very weak, similarities between species. Effector proteins are those which present the best biological, vaccine and diagnostic potentialities.

The inventors have discovered that the clustering of class I and class II genes inside a single pathogenicity island, offers the opportunity of conveniently finding and characterising unknown class II genes by targeting class I genes which can be identified using a known sequence of one of their numerous orthologues.

Bordetella pertussis

Whooping cough is a disease caused by infection by *Bordetella pertussis*, and is a serious and debilitating human disease particularly in young children. Although whole cell and acellular vaccines are available that are effective against the disease, there remains a need for the identification of further highly purified pertussis proteins that could be used in a more efficacious pertussis vaccine.

Although many pertussis virulence associated factors are known such as pertussis toxin, filamentous haemagglutinin, pertactin, which have been included in various acellular vaccines, there is no convenient genetic method for identifying further virulence factors using the pertussis genome (short of laboriously sequencing the whole genome). Although class I type III secretion system virulence genes have recently been shown to exist in *B. bronchiseptica* and *B. pertussis* (Yuk *et al.*, 1998), there has been no complete analysis of a pathogenicity island in *Bordetella*, and the identity and characterisation of effector genes within such a pathogenicity island has been unknown up until the present invention.

SUMMARY OF THE INVENTION

In one aspect, the invention relates to a method for the identification of new virulence genes in bacterial strains containing a type III secretion system. In particular, the invention allows the identification of the effector virulence genes associated within a pathogenicity island containing the genes for the type III secretion system. Another aspect of the invention a method for the identification of pathogenic bacterial strains containing a type III secretion system. Another aspect of the invention relates to *Bordetella pertussis* BopN, Orf1, Orf2, Orf3, Orf4, Orf5, Orf6, Orf7, Orf8, Orf9, Orf10, Orf11, Orf12, Orf13, Orf14, Orf15 effector proteins, and the respective polynucleotide sequences encoding them.

Although the general concepts of type III secretion systems and pathogenicity islands have been reported, the problem of how simply and reliably to identify whether any given organism has such cell machinery has not been accomplished until now. Such a method is extremely useful to establish whether a given strain has a type III secretion system within a pathogenicity island, to characterise unknown virulence genes within the pathogenicity island, and to use in quick diagnostic methods for determining whether a cultured bacterial strain containing a type III secretion system is pathogenic.

In the present invention, a novel, general method is described to achieve the above aims. More specifically, the invention utilises a method that employs ideally-suited primers designed specifically from the sequence of the virulent *Yersinia enterocolitica lcrD* gene as a target sequence. The presence of a type III secretion system within a pathogenicity island in *Bordetella pertussis* was discovered, and every gene within the pathogenicity island was characterised.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1. Nucleotide and deduced amino acid sequences of the cloned 152 bp amplicon. The primers involved in the original amplification, the subsequent nested PCR, and the gene library screening are all derived from this sequence, and listed specifically in Table 1.

Fig. 2. PileUp figure from the deduced amino acid sequences homologous to *Yersinia* LcrD. Abbreviations: BbuFlhA = *Borrelia burgdorferi* FlhA; TpaFlhA = *Treponema pallidum* FlhA; BsuFlhA = *Bacillus subtilis* FlhA; CjeFlbA = *Campylobacter jejuni* FlbA; HpyFlhA = *Helicobacter pylori* FlhA; EcoFlhA - *Escherichia coli* FlhA; StyFlhA = *Salmonella typhimurium* FlhA; YenFlhA = *Yersinia enterocolitica* FlhA; PmiFlhA = *Proteus mirabilis* FlhA; CcrFlbF = *Caulobacter crescentus* FlbF; EcoFhiA = *Escherichia coli* FhiA; EamHrpI = *Erwinia amylovora* HrpI; PsyHrpI = *Pseudomonas syringae* HrpI; ECEPSepA = Enteropathogenic *Escherichia coli* SepA; StySsaV = *Salmonella typhimurium* SsaV; RsoHrpO = *Ralstonia solanacearum* HrpO; XcaHrpC2 = *Xanthomonas campestris* HrpC2; SflMxiA = *Shigella Flexneri* MxiA; StyInvA = *Salmonella typhimurium* InvA; PaePcrD = *Pseudomonas aeruginosa* PcrD; YenLcrD = *Yersinia enterocolitica* LcrD; BpeBcrD = *Bordetella pertussis* BcrD; CpsTtsB = *Chlamydia psittaci* TtsB.

Fig. 3. Organization of the *Bordetella pertussis* pathogenicity island (Pai). Four house keeping genes (hatched boxes) and the transposase gene of IS*481* (black box) are surrounding the Pai. The Pai consists of genes coding for determinants involved in the secretory apparatus and its regulation (class I genes, in grey boxes) as well as ORFs which putitively code for effector proteins (class II genes, in white boxes). Letters indicate the respective class I *bsc* genes whereas numbers correspond to the class II ORFs listed in Table 3.

Fig. 4. PileUp figure from the deduced amino acid sequences homologous to *Yersinia* YscU. Abbreviations: BbuFlhB = *Borrelia burgdorferi* FlhB; TpaFlhB = *Treponema pallidum* FlhB; EcoFlhB - *Escherichia coli* FlhB; StyFlhB = *Salmonella typhimurium* FlhB; PmiFlhBpart = partial *Proteus mirabilis* FlhB; YenFlhB = *Yersinia enterocolitica* FlhB; BsuFlhB = *Bacillus subtilis* FlhB; HpyFlhB = *Helicobacter pylori* FlhB; AtuFlhB = *Agrobacterium tumefaciens* FlhB; CcrPodW = *Caulobacter crescentus* PodW; SflSpa40 = *Shigella flexneri* Spa40; StySpaS = *Salmonella typhimurium* SpaS; EcoEscU = *Escherichia coli* EscU; StySsaU = *Salmonella typhimurium* SsaU; BpeBscU = *Bordetella pertussis* BscU; YenYscU = *Yersinia enterocolitica* YscU; RsoHrpN = *Ralstonia solanacearum* HrpN; XcaOrf0part = partial *Xanthomonas campestris* Orf0; EamHrcU = *Erwinia amylovora* HrcU; EheHrcUpart = partial *Erwinia herbicola* HrcU; PsyHrpY = *Pseudomonas syringae* HrpY; CpsOrf1 = *Chlamydia psittaci* Orf1.

Fig. 5. The DNA sequence of the Bordatella pertussis genome comprising the type III secretion system pathogenicity island. Reference should be made to tables 2, 3, and 4 and Fig. 3 for information regarding open reading frames.

Fig. 6. Purification of MBP-Orf2, -4, -6 and –10 by affinity chromatography. The ultracentrifugatrion supernatants of each lysate (left part of the panels) and the products eluated from the affinity column (right part of the panels) were analysed by SDS-PAGE and revealed by Coomassie blue staining.

DESCRIPTION OF THE INVENTION

Type III secretion systems identified to date are encoded by either chromosal or plasmidic pathogenicity island genes. However, no where in the prior art was it realised that the conservation of genes encoding class I components of type III secretion systems and the clustering of these genes with effector protein coding sequences offered the opportunity for detecting unidentified target proteins involved in host colonisation. Such proteins would be potentially valuable in both vaccinal and diagnostic fields.

Although the known sequence of a gene encoding any conserved (class I) type III secretion machinery protein can be used in performing this invention, the *lcrD* gene is preferred. The chosen gene will act as a target for detecting unidentified pathogenicity islands in related bacterial species. The *lcrD* gene from *Yersinia* is preferred as it codes for the archetype of the recently identified LcrD/FlbF family of proteins. Members of this family are involved in host cell invasion, virulence in several phytopathogenic bacteria or in flagellar assembly. *lcrD* is preferred because the LcrD protein, and consequently the gene encoding it, is one of the most conserved determinants of the secretion machinary. Additionally, multiple amino acid comparisons have shown that the classification of the LcrD family members can be split into two main subfamilies, which, interestingly, can be correlated with the functions assigned to these proteins of each subfamily. One subfamily encompasses all the motility-involved proteins, while the other encompasses all the virulence-related determinants. This observation is illustrated in Fig. 2 (and mentioned in Gyri *et al.* (1995) & Bogdanove *et al.* (1996)). Thus, if an unknown *lcrD* homologous gene is identified, it may, after being routinely sequenced, be classified as a virulence or a flagellar gene. Once the pathogenicity island is identified, this simple test would therefore define whether the search for other virulence genes on the pathogenicity island should be initiated.

The preferred method for identifying unknown pathogenicity islands comprising a type III secretion system is by:

i) identifying two highly conserved regions of the target protein sequence (preferably of LcrD). Preferably, both regions should contain conserved amino acids which are encoded by the fewest number of codon possibilities e.g. Methionine (ATG being the only possibility) or Tryptophan (TGG being the only possibility). This minimises the number of permutations in both degenerate primer sets that are designed in the next stage of the process, thus ensuring a greater probability that each primer set will specifically anneal to the unknown *lcrD*-equivalent gene (thereby minimising background non-specific interactions). Most preferably, regions should also be chosen that are clearly distinguishable from the paralogue *flhA* flagellar genes, present in all flagellated bacterial strains.

ii) designing a degenerate set of primers for both of the chosen regions such that a) the primers are at least 15 bases long, preferably 20-30 bases long, and still more preferably 21-23 bases long, b) they are degenerate at bases that can be more than one type of nucleotide whilst still encoding the same amino acid (due to the degeneracy of codon usage for amino acids), but no more degenerate than is required to cover all permutations for the amino acid region selected, and c) the primer set that encodes the more N-terminal region of the chosen protein should correspond to the coding strand of its corresponding double-stranded DNA sequence, and the set that encodes the more C-terminal region should correspond to the complementary strand of the corresponding double-stranded DNA sequence.

iii) synthesising the degenerate primer sets of step ii) using conventional DNA synthesis methods well known in the art.

iv) purifying the primer sets of step iii)

v) adding both the primer sets and a sample containing nucleic acid from a bacterial strain (preferably a cell sample of the bacterial species itself) together in appropriate quantities and in an appropriate buffer in order to perform a polymerase chain reaction (PCR)

vi) performing a PCR reaction in order to amplify the region of the gene between the two primers (conditions for performing the PCR reaction can be optimised using techniques well known in the art)

vii) observing the reaction products on a gel (preferably an agarose gel) for an amplified product of the size expected; if no such product is present, the bacterial strain is unlikely to use a type III secretion system; if such a product is present, the bacterial strain is likely to have a type III secretion system, and is likely to be pathogenic.

The preferred method for confirming that the amplified product actually corresponds to a virulence gene is by carrying out steps i)-vii) above (where the target protein is LcrD) and then:

viii) optionally separating the product of correct size from any background products of incorrect size by removing the correct band from the gel, purifying the product by conventional means, and amplifying the product once more with the two degenerate primer sets in another PCR reaction (under preferably more stringent PCR conditions) [this step is required should the product of step vii) not be pure enough for direct cloning]

ix) inserting the DNA fragment by conventional means into a vector which is capable of being sequenced, and sequencing the fragment

x) comparing the deduced amino acid sequence of ix) with that of known members of the LcrD/FlbF family of proteins to associate the amplified product as being part of either a virulence or a flagellar gene.

And optionally:

xi) using the internal sequence of the fragment to design primers that are the exact sequence of, and specific to, the unknown *lcrD*-equivalent gene.

xii) using the primers of xi) firstly to screen a genomic library of the organism for positive clones

xiii) isolating the clones of xii), and sequence one or more of said clones

xiv) scanning the sequence of one clone (and overlapping sequences of other clones) to search for an open reading frame which is approximately the same size as *lcrD* (approximately 2100bp), and encodes a protein homologous to LcrD

xv) ascertaining whether the LcrD-equivalent protein is more homologous with the *flbF* (flagellar protein secretion) gene family or the *lcrD* (type III secretion system pathogenicity island) gene family.

The preferred method for characterising the whole pathogenicity island and defining unidentified virulence effector genes is by carrying out steps i)-xv) above (where the target protein is LcrD) and then:

xvi) if the sequence is more homologous with the *lcrD* gene family, designing primers at either extreme of the gene sequence already ascertained, and scanning and sequencing the genomic library (using a standard chromosome walking strategy – where the insert boundaries of an original clone serves as a probe for screening and cloning adjacent regions) to sequence eventually the whole of the pathogenicity island (both boundaries of which will be defined by the presence of either direct or inverted repeats, or insertion sequences, or the presence of house-keeping genes)

xvii)defining unidentified virulence effector genes within the sequenced pathogenicity island

xviii)cloning, expressing and characterising the virulence genes of xvii) which encode virulence effector proteins of the organism

Definitions

"Bordetella pathogenicity proteins" refers generally to polypeptides having the amino acid sequence encoded by the genes defined in tables 2 and 3, or an allelic variant thereof. These proteins are: BcrD, BcrH, BscC, BscD, BscE, BscF, BscI, BscJ, BscK, BscL, BscN, BscO, BscP, BscQ, BscR, BscS, BscT, BscU, BscV, BrpL, BopN, Orf1, Orf2, Orf3, Orf4, Orf5, Orf6, Orf7, Orf8, Orf9, Orf10, Orf11, Orf12, Orf13, Orf14, Orf15.

"Bordetella pathogenicity genes" refers to polynucleotides having the nucleotide sequence defined in tables 2 and 3, or allelic variants thereof and/or their complements. These genes are: *bcrD, bcrH, bscC, bscD, bscE, bscF, bscI, bscJ, bscK, bscL, bscN, bscO, bscP, bscQ, bscR, bscS, bscT, bscU, bscV, brpL, bopN, orf1, orf2, orf3, orf4, orf5, orf6, orf7, orf8, orf9, orf10, orf11, orf12, orf13, orf14, orf15*.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslational natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation,

proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Posttranslational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential

properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions (preferably conservative), additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis. Variants should retain one or more of the biological activities of the reference polypeptide. For instance, they should have similar (preferably the same) antigenic or immunogenic activities as the reference polypeptide. Antigenicity can be tested using standard immunoblot experiments, preferably using polyclonal sera against the reference polypeptide. The immunogenicity can best be tested by measuring antibody responses (using polyclonal sera generated against the variant polypeptide) against purified reference polypeptide in a standard ELISA test. Preferably, a variant would retain all of the above biological activities.

"Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New

York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heijne, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., *et al.*, *Nucleic Acids Research* (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. *et al.*, *J Molec Biol* (1990) 215:403). Most preferably, the program used to determine identity levels was the GCG 9 package, as was used in the Examples below.

As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include on average up to five point mutations per each 100 nucleotides of the reference nucleotide sequence. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal

positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

Polypeptides of the invention

In one aspect, the present invention relates to Bordetella pathogenicity proteins (or polypeptides). The Bordetella pathogenicity polypeptides include the polypeptides encoded by the genes defined in tables 2 and 3; as well as polypeptides comprising the amino acid sequence encoded by the genes defined in tables 2 and 3; and polypeptides comprising the amino acid sequence which have at least 75% identity to that encoded by the genes defined in tables 2 and 3 over their entire length, and preferably at least 80% identity, and more preferably at least 90% identity. Those with 95-99% identity are highly preferred.

The Bordetella pathogenicity polypeptides (or fragments thereof) may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It may be advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues or Maltose Binding Protein (MBP), or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

Fragments of the Bordetella pathogenicity polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that is the same as part, but not all, of the amino acid sequence of the aforementioned Bordetella pathogenicity polypeptides. As with Bordetella pathogenicity polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number

1-20, 21-40, 41-60, 61-80, 81-100, and 101 to the end of Bordetella pathogenicity polypeptide. In this context "about" includes the particularly recited ranges larger or smaller by several, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes. The fragments should comprise at least 7 consecutive amino acids from the sequences e.g. 8, 10, 12, 14, 18, 20 or more depending on the particular sequence). Preferably the fragments comprise an epitope from the sequence.

Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of Bordetella pathogenicity polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus and/or transmembrane region or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Other preferred fragments are biologically active fragments. Biologically active fragments are those that mediate Bordetella pathogenicity protein activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

Preferably, all of these polypeptide fragments retain the biological activity (for instance antigenic or immunogenic) of the Bordetella pathogenicity protein, including antigenic activity. Variants of the defined sequence and fragments also form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic

residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination. Most preferred variants are naturally occurring allelic variants of Bordetella pathogenicity polypeptide present in strains of *Bordetella pertussis*.

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

The Bordetella pathogenicity polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

It is most preferred that a polypeptide of the invention is derived from *Bordetella pertussis*, however, it may preferably be obtained from other organisms of the same taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order, such as *Bordetella parapertussis* or *Bordetella bronchiseptica*.

A further aspect of the invention is substantially purified Bordetella pathogenicity polypeptides of the invention. "substantially purified" when used in reference to a protein or peptide means that the molecule has been largely, but not necessarily wholly, separated an purified from other cellular and non-cellular components. Typically a protein is substantially pure when it is at least about 60 % by weight free from other

naturally occurring organic molecules. Preferably the purity is at least about 75 %, more preferably at least about 90% , and most preferably at least about 99% by weight pure.

Polynucleotides of the invention

Another aspect of the invention relates to Bordetella pathogenicity polynucleotides. Bordetella pathogenicity polynucleotides include isolated polynucleotides which encode the Bordetella pathogenicity polypeptides and fragments respectively, and polynucleotides closely related thereto or variants thereof. More specifically, Bordetella pathogenicity polynucleotides of the invention include a polynucleotide comprising the nucleotide sequence of genes defined in table 2 or 3, encoding a Bordetella pathogenicity polypeptide. Bordetella pathogenicity polynucleotides further include a polynucleotide comprising a nucleotide sequence that has at least 75% identity over its entire length to a nucleotide sequence encoding the Bordetella pathogenicity polypeptide encoded by the genes defined in tables 2 and 3, and a polynucleotide comprising a nucleotide sequence that is at least 75% identical to that of the genes defined in tables 2 and 3. In this regard, polynucleotides at least 80% identical are particularly preferred, and those with at least 90% are especially preferred. Furthermore, those with at least 95% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. Also included under Bordetella pathogenicity polynucleotides is a nucleotide sequence which has sufficient identity to a nucleotide sequence of a gene defined in tables 2 and 3 to hybridize under conditions useable for amplification or for use as a probe or marker. The invention also provides polynucleotides which are complementary to such Bordetella pathogenicity polynucleotides.

Using the information provided herein, such as specific Bordetella pathogenicity gene and polypeptide sequences, a polynucleotide of the invention encoding a Bordetella pathogenicity polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *Bordetella pertussis* cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, typically a

library of clones of chromosomal DNA of *Bordetella pertussis* in *E.coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent hybridization conditions. By sequencing the individual clones thus identified by hybridization with sequencing primers designed from the original polypeptide or polynucleotide sequence it is then possible to extend the polynucleotide sequence in both directions to determine a full length gene sequence. Conveniently, such sequencing is performed, for example, using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Direct genomic DNA sequencing may also be performed to obtain a full length gene sequence.

A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than *Bordetella pertussis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions (for example, using a temperature in the range of 45 – 65°C and an SDS concentration from 0.1 – 1%) with a labeled or detectable probe consisting of or comprising a sequence defined in table 2 or 3 or a fragment thereof; and isolating a full-length gene and/or genomic clones containing said polynucleotide sequence.

The invention also provides a polynucleotide consisting of or comprising a polynucleotide sequence obtained by screening an appropriate library containing the complete gene for a polynucleotide sequence defined in tables 2 and 3 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence defined in table 2 or 3 or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers are described elsewhere herein.

The nucleotide sequence encoding Bordetella pathogenicity polypeptide encoded by the genes defined in tables 2 and 3 may be identical to the polypeptide encoding sequence contained in the genes defined in tables 2 or 3, or it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide encoded by the genes defined in tables 2 and 3 respectively.

When the polynucleotides of the invention are used for the recombinant production of Bordetella pathogenicity polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro-protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag, or is glutathione-s-transferase, or is MBP. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Nucleic acid comprising fragments of the sequences of the invention are also provided. These should comprise at least 10 consecutive nucleotides from the sequences (e.g. 12, 14, 15, 18, 20, 25, 30, 35, 40 or more depending on the particular sequence). Such fragments can preferably hybridise to the above-mentioned sequences under stringent conditions.

Further preferred embodiments are polynucleotides encoding Bordetella pathogenicity protein variants comprising the amino acid sequence of the Bordetella pathogenicity polypeptide encoded by the genes defined by tables 2 and 3 respectively in

which several, 10-25, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination. Most preferred variant polynucleotides are those naturally occurring *Bordetella pertussis* sequences that encode allelic variants of the Bordetella pathogenicity proteins in *Bordetella* strains, preferably *B. pertussis*.

The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 80%, and preferably at least 90%, and more preferably at least 95%, yet even more preferably 97-99% identity between the sequences.

Polynucleotides of the invention, which are identical or sufficiently identical to a nucleotide sequence of any gene defined in tables 2 and 3 or a fragment thereof, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding Bordetella pathogenicity polypeptides respectively and to isolate cDNA and genomic clones of other genes (including genes encoding homologs and orthologs from species other than *Bordetella pertussis*) that have a high sequence similarity to the Bordetella pathogenicity genes. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide sequences are 80% identical, preferably 90% identical, more preferably 95% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides. In one embodiment, to obtain a polynucleotide encoding Bordetella pathogenicity polypeptide, including homologs and orthologs from species other than *Bordetella pertussis*, comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having a nucleotide sequence contained in one of the gene sequences defined by tables 2 and 3, or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Thus in another aspect, Bordetella pathogenicity polynucleotides of the present invention further

include a nucleotide sequence comprising a nucleotide sequence that hybridize under stringent condition to a nucleotide sequence having a nucleotide sequence contained in one of the genes defined by table 2 and 3, or a fragment thereof. Also included with Bordetella pathogenicity polypeptides are polypeptides comprising amino acid sequences encoded by nucleotide sequences obtained by the above hybridization conditions. Such hybridization techniques are well known to those of skill in the art. Stringent hybridization conditions are as defined above or, alternatively, conditions under overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

A coding region of a Bordetella pathogenicity gene may be isolated by screening using a DNA sequence defined in table 2 or 3 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman, *et al.*, *PNAS USA 85:* 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5'

in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the product directly to the existing DNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

The polynucleotides of the invention that are oligonucleotides derived from a sequence defined in table 2 or 3 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

Diagnostic Assays

This invention also relates to the use of Bordetella pathogenicity polypeptides, or Bordetella pathogenicity polynucleotides, for use as diagnostic reagents. Detection of Bordetella pathogenicity polypeptides will provide a diagnostic tool that can add to or define a diagnosis of *B. pertussis* disease, among others.

Materials for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy.

Thus in another aspect, the present invention relates to a diagonostic kit for a disease or suspectability to a disease, particularly *B. pertussis* disease, which comprises:
(a) a Bordetella pathogenicity polynucleotide, preferably the nucleotide sequence of one of the gene sequences defined by tables 2 and 3, or a fragment thereof;
(b) a nucleotide sequence complementary to that of (a);

(c) a Bordetella pathogenicity polypeptide, preferably the polypeptide encoded by one of the gene sequences defined in tables 2 and 3, or a fragment thereof;
(d) an antibody to a Bordetella pathogenicity polypeptide, preferably to the polypeptide encoded by one of the gene sequences defined in tables 2 and 3; or
(e) a phage displaying an antibody to a Bordetella pathogenicity polypeptide, preferably to the polypeptide encoded by one of the gene sequences defined in tables 2 and 3.

It will be appreciated that in any such kit, (a), (b), (c), (d) or (e) may comprise a substantial component.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence selected from a related organism, preferably a different species of the same genus or a different strain of the same species. Point mutations can be identified by hybridizing amplified DNA to labeled Bordetella pathogenicity polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al.*, *Science, 230:* 1242 (1985). Sequence changes at specific locations also may be

revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA, 85:* 4397-4401 (1985).

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, which is associated with a disease or pathogenicity will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such as those described elsewhere herein.

The invention further provides a process for diagnosing disease, preferably bacterial (particularly *Bordetella*) infections, more preferably infections caused by *Bordetella pertussis*, comprising determining from a sample derived from an individual, such as a bodily material, an increased level of expression of polynucleotide having a sequence defined in table 2 or 3. Increased or decreased expression of a polynucleotide can be measured using any on of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting, spectrometry and other hybridization methods.

Vectors, Host Cells, Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems that comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of the polypeptides of the invention, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci, *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, *Moraxella catarrhalis*, *Haemophilus influenzae* and *Neisseria meningitidis*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, a basidiomycete, *Candida albicans* and *Aspergillus*; insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia

viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses, retroviruses, and alphaviruses and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia,

fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelian Equine Encephalitis Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus), herpesviruses (varicella zoster virus, etc), Listeria, Salmonella , Shigella, Neisseria, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Antibodies

According to a further aspect, the invention provides antibodies which bind specifically to the polypeptides of the invention. These may be polyclonal or monoclonal and may be produced by any suitable means well known to a skilled person in the art.

Typically, a mouse or rat is immunised with a protein (preferably adjuvanted with Freund's complete adjuvant) and injected (doses of 50-200 μg/injection is typically sufficient). Polyclonal antibodies can be isolated by bleeding the animal to extract serum. Alternatively, monoclonal antibodies can be generated by removing the spleen (or large lymph nodes) and dissociating it into single cells (Kohler and Milstein, (1975) Nature, 256:495-497). These are then induced to fuse with myeloma cells to form hybridoma, and are cultured in a selective medium (eg hypoxanthine, aminopterin, thymidine merium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected monoclonal-secreting hybridomas are then cultured either in vitro (eg in tissue culture bottles or hollow fiber reactors), or in vivo (as Ascites in mice).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms or animals, such as other mammals, may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-Bordetella pathogenicity polypeptide or from naive libraries (McCafferty, *et al.*, (1990), Nature 348, 552-554; Marks, *et al.*, (1992) *Biotechnology 10*, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature 352:* 628).

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

Antibodies against a Bordetella pathogenicity polypeptide or polynucleotide may be employed to treat infections, particularly bacterial infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complimentarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with Bordetella pathogenicity polypeptide or epitope-bearing fragments, analogs, outer-

membrane vesicles or cells (attenuated or otherwise) adequate to produce antibody and/or T cell immune response to protect said animal from Bordetella (particularly *B. pertussis*) disease, among others. Such agents may be used alone, or conjugated to another molecule which improves its immunological potency. In particular the invention relates to the use of Bordetella pathogenicity polypeptides encoded by the genes defined in table 3 – the effector proteins. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering Bordetella pathogenicity polypeptide via a vector directing expression of Bordetella pathogenicity polynucleotide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

A further aspect of the invention relates to an immunological composition or vaccine formulation which, when introduced into a mammalian host, induces an immunological response in that mammal to a Bordetella pathogenicity polypeptide (particularly one encoded by a gene defined in table 3) wherein the composition comprises a Bordetella pathogenicity gene, or Bordetella pathogenicity polypeptide or epitope-bearing fragments, analogs, outer-membrane vesicles or cells (attenuated or otherwise). The vaccine formulation may further comprise a suitable carrier. The Bordetella pathogenicity polypeptide vaccine composition is preferably administered orally or parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems

known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

The vaccine formulations of the invention may also comprise other *Bordetella* antigens known to be suitable vaccinal agents, for instance: pertussis toxoid, pertactin, agglutinogins 1 and 2, FHA (filamentous haemagglutinin), and adenylate cyclase / haemolysin (AC/HLY), or immunogenic fragments thereof (Locht *et al.*, NAR (1986) 14:3251-3261; Relman *et al.*, PNAS USA (1989) 86:2637-2641; Roberts *et al.*, Mol. Microbiol. (1991) 5:1393-1404; Mooi *et al.*, Microb. Pathog. (1992) 12:127-135; Hewlett and Gordon, In *Pathogenesis and Immunity in Pertussis* (1988), New York, Wiley & Sons, pp. 193-209.

Yet another aspect of the invention relates to an immunological/vaccine formulation which comprises the polynucleotide of the invention. Such techniques are known in the art, see for example Wolff *et al.*, *Science*, (1990) *247*: 1465-8.

Vaccine compositions can comprise polypeptides, antibodies, or polynucleotides of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition (in this case *Bordetella*, particularly *B. pertussis*, disease), or to exhibit a detectable therepeutic or preventative effect. The effect can be detected by, for example, antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature. Immunogenic compositions used as vaccines comprise an immunologically effective amount of the antigenic or immunogenic polypeptides. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention.

EXAMPLES

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples illustrate, but do not limit the invention.

Example 1: A type III secretion system is present in a pathogenicity island in *Bordetella pertussis*.

The presence of a *lcrD* homologous gene in the *Bordetella pertussis* genome was investigated by polymerase chain reaction (PCR). The primers used (oligos 95080 and 95081 shown in Table 1) were degenerate oligonucleotides corresponding to highly conserved regions of the amino acids sequences of the LcrD/FlbF family of proteins. These primers were also designed to favour the amplification of virulence genes instead of their paralogue *flhA* or *flbF* flagellar genes, present in flagellated bacterial strains. The presence of the 3' triplet CAT in oligonucleotide 95081 is a determinant – indeed when multiple sequence analysis is done using known homologous sequences (database searching was done with either the FASTA and TFASTA programs of the GCG9 package, or with BLASTN, BLASTP and BLASTX programs, and alignments were carried out with the PILEUP program from the GCG9 package) it could be seen that the CAT triplet codes for a methionine which is exclusively present in virulence sequences while absent in the flagellar ones.

When analysed on agarose gel, the PCR product appeared as a heterogeneous mix of fragments, one of which was presenting the expected size (around 150 bp). A second round of amplification using the approximately 150 bp DNA as template yielded a single amplicon which was cloned in pCRII (obtained from Invitrogen) for further characterisation. It appeared as a 152 bp fragment whose nucleotide sequence (Fig. 1), although similar to all *lcrD/flbF* homologous genes, shares a higher level of identity with the virulence (*lcrD*-like) genes.

Table 1.

oligonucleotides	sequence[1]	features	*lcrD* corresponding codons[2]
95080	GSH ATG CCW GGH AAR CAR ATG	direct, degenerate	150 to 156
95081	GC RTC DCC YTT DAC RAA YTT CAT	complement, degenerate	193 to 200
95363	CC ATC GAC GCG GAC TTG CGC G	direct, non-degenerate	157 to 164
95364	CGC GCC GTC CAT GGC GCC ATA	complement, non-degenerate	186 to 192
96110	C CGA CGC CGA CGC CGT ACG GTC	direct, non-degenerate	172 to 179

[1] The letter code for nucleotide ambiguity proposed by IUB (Nomenclature Committee, 1985, Eur. J. Biochem., **150**: 1-5) was used.

[2] The DNA sequence of the *lcrD* gene from *Yersinia enterocolitica* used for this work was published by Plano *et al.* (1991).

To ensure that the cloned fragment was actually a *B. pertussis* sequence PCR was performed under stringent conditions with serial 10-fold dilutions of DNA from *B. pertussis*. The optimisation of stringent PCR conditions require a perfect match between template and primers. It was likely, however, that due to the degeneration of the original primers, the 152 bp sequence initially obtained had, at its boundaries, a few base pair differences with the actual *B. pertussis lcrD*-like (hereafter called *bcrD*) sequence. A nested PCR approach using internal primers (oligos 95363 and 95364 Table 1) was therefore preferred, as primers known to be the correct *B. pertussis* sequence are used. A dose-response-relationship was observed between the 10-fold dilutions of *B. pertussis* template DNA and the product of the nested PCR, suggesting that the 152 bp amplicon actually originates from the *Bordetella* genome.

Comparison of the 152 bp sequence with *lcrD/flbF* genes allowed us to define a specific DNA stretch (oligo 96110 in Table 1) which was used as a probe for screening a

genomic library of *B. pertussis* constructed in the plasmid vector pBR327 (Delisse-Gathoye *et al*, 1990, *Infect-Immun.* **58**: 2895-905). Several positive clones were isolated and restriction analysis of their resident plasmids showed that they harboured overlapping inserts. The entire nucleotide sequence of one insert was determined, revealing a large open reading frame (ORF). This 2100 bp ORF encoded a 75 kDa polypeptide which is 59 % and 47 % identical to the yersinial proteins LcrD and FlhA respectively. Multiple amino acids comparisons of all known members of the LcrD/FlbF family of proteins, including the *B. pertussis* BcrD deduced amino acid sequence, showed that this sequence clearly ranked within the virulence associated determinants (Fig. 2). These data strongly suggest that *B. pertussis* possesses a type III export system, involved in the secretion of virulence effectors.

The *B. pertussis lcrD*-like nucleotide sequence (*bcrD*) has been submitted to EMBL and assigned the accession number Y13383.

This general technique has been useful for determining the presence/absence of a type III secretion system in other bacterial strains. The human pathogens *Borrelia burgdorferi* and *Helicobacter pylori* were intensively screened for such a system using this technique. No evidence for a type III secretion system could be found. The subsequent publication of the genome sequences of these microorganisms has confirmed the absence of similar systems in these species. In contrast, the method allowed the amplification of a DNA fragment from the phytopathogen *Pseudomonas corrugata*, which clearly ranks among the virulence sequences. This technique could be applied to any Gram negative pathogen of medical or agronomic importance such as *Neisseria* spp, *Moraxella catharalis*, *Vibrio cholerae*, any Enterobacteriaceae, *Pseudomonas* spp, *Haemophilus influenzae*, *Brucella* spp, *Francisella tularensis*, *Pasteurella* spp, *Legionella pneumophila*. Even in strains that have been fully sequenced, this technique can be used as a simple method for checking alternate types or strains of the same species. For instance, some types of pathogenic *Escherichia coli* harbour a type III secretion system whereas others do not.

Example 2: Analysis of the *B. pertussis bcrD* flanking sequences to characterise the pathogenicity island and virulence-related proteins encoded therein

The tendency for systematic clustering of type III encoding genes inside pathogenicity islands prompted the analysis of *B. pertussis bcrD* flanking sequences. The whole region containing the pathogenicity island was sequenced by chromosome walking taking care to pay attention to the fact that each Pathogenicity island region must be represented in at least two independent clones, to avoid possible artefacts due to chimeric DNA inserts. This revealed clustered ORFs that could be classed in 3 categories: class I type ORFs (table 2); class II type ORFs (table 3) – the effector proteins which have the best vaccinal and diagnostic properties; & insertion sequences, and ORFs homologous to house keeping genes of other species (table 4). Although there is no general rule for defining the boundaries of a Pathogenicity island, they can be demarcated with a direct or inverse repeat at one or other boundary, however the absolute demarcation of the boundaries can only really be done by the detection of house keeping genes at the extremes of the sequence. In the present case, an insertion sequence (IS in Fig. 3) was present at the 5' end of the island (separating the virulence ORFs from the house keeping genes), but absent at the 3' end. In addition, the presence of house keeping genes (*greA* and ICFG-like) surrounding a locus which, according to sequence data, encompasses numerous virulence sequences is a good indication of the boundaries of the island. The complete gene organisation of the pathogenicity island is schematically represented in Figure 3. The precise definition of the PAI boundaries requires further experimental data, such as the characterisation of the corresponding chromosomal region of a *Bordetella* strain which is devoid of a type III secretion system.

Table 2

names	Coding sequence from/to (with reference to Fig. 5)	Coding DNA strand	SEQ ID NO:	Homologous genes (from *Yersinia*, unless otherwise specified)
Class I genes, i.e. genes coding for determinants involved in the secretory apparatus and their regulation				
bcrD	8656/10755	complement	1	*LcrD*
bcrH	14097/14582	direct	3	*lcrH* (= *sycD*)
bscC	26955/28757	direct	5	*YscC*
bscD	7379/8659	complement	7	*YscD*
bscE	7039/7338	complement	9	None
bscF	6783/7049	complement	11	*YscF*
bscI	17892/18218	direct	13	*YscI*
bscJ	18215/19039	direct	15	*YscJ*
bscK	19032/19694	direct	17	None
bscL	19664/20302	direct	19	*YscL*
bscN	20307/21641	direct	21	*YscN*
bscO	21641/22150	direct	23	*YscO*
bscP	22147/22695	direct	25	None
bscQ	22692/23771	direct	27	*YscQ*
bscR	23768/24439	direct	29	*YscR*
bscS	24445/24711	direct	31	*YscS*
bscT	24723/25523	direct	33	*YscT*
bscU	25520/26569	direct	35	*YscU*
bscV	26566/26964	direct	37	None
brpL	28778/29380	complement	39	*hrpL* (*Pseudomonas syringae*)

Table 3

Names	Coding sequence from/to (with reference to Fig. 5)	Coding DNA strand	SEQ ID NO:	Homologous genes (from *Yersinia*, unless otherwise specified)
Class II ORFs which putatively code for effector proteins				
bopN	11906/13003	complement	41	*YopN* (= *lcrE*)
orf1	6160/6747	direct	43	None
orf2	10752/11120	complement	45	None
orf3	11117/11527	complement	47	None
orf4	11532/11909	complement	49	None
orf5	13002/13784	direct	51	None
orf6	13806/14081	direct	53	None
orf7	14630/15571	direct	55	None
orf8	15601/16803	direct	57	None
orf9	16827/17288	direct	59	*BcrH*
orf10	17293/17814	direct	61	*pcr4* (*Pseudomonas aeruginosa*)
orf11	29412/29591	complement	63	None
orf12	29555/30529	complement	65	None
orf13	30631/31776	direct	67	None
orf14	31773/33005	complement	69	None
orf15	32370/33014	direct	71	None

Table 4

No name specified	Coding sequence from/to (with reference to Fig. 5)	Coding DNA strand	SEQ ID NO:	Homolgous sequences
Insertion Sequences and house keeping genes				
	711/2024	direct	73	uracil permease genes of numerous bacteria
	2055/3590	complement	75	Chemoreceptor genes of numerous bacteria
	4220/4696	direct	77	*greA* (*Escherichia coli*)
	4998/5948	complement	79	transposase genes of numerous bacteria
	33002/34852	complement	81	ICFG gene (*Synechocystis* sp)

Next to the *bcrD* gene, there is an open reading frame (ORF) whose deduced amino acid sequence shares significant similarities with the YscU protein of *Yersinia* spp (39% identity and 51% similarity) and other known YscU homologs (Fig. 4). YscU, like LcrD, is a component of the *Yersinia* type III secretion machinery involved in the virulence mechanisms of the bacteria. *B. pertussis* therefore possesses a classical type III secretion system which is most probably involved in pathogenicity. This latter point can be investigated through phentoypic analyses of mutants (see below).

The total length of the Pai is approximately 30 to 40 kb. The DNA sequence of the whole region is presented in Figure 5, and is referred to in tables 2, 3, and 4. Restriction analysis on pulsed-field gel electrophoresis allowed the type III locus to be mapped at coordinate position 1,590 kb on the Tohama I strain chromosome.

No homologies could be found between the *B. pertussis* Class II Pai DNA sequences and the sequences reported in the GenEMBL databases (except for those stated in table 3). The expressed products of these unknown genes within the Pai

responsible for virulence, will be useful in the development of a vaccine formulation against pathogenic *Bordetella pertussis*.

To address the precise function of the Pai, a *bcrD* mutant was engineered by allelic exchange. In the resulting mutant, the *bcrD* gene was disrupted by an *aphA*-3 cassette conferring kanamycin resistance. This cassette was inserted in such a way that translation was not interrupted, avoiding any polar effect on expression of putative downstream cistrons. A mutant has been isolated and its associated phenotype is being currently analysed.

Example 3: Analysis of the *in situ* expression of the genes of the pathogenicity island

Genetic constructions

To produce a mutant defective in type III secretion, a 255-bp fragment (codons 363 to 445) was deleted from the *bcrD* coding sequence and replaced by a cassette containing the *aphA-3* gene which confers kanamycin resistance (Menard *et al.*, J. Bacteriol. (1993) 175:5899-5906). The *aphA-3* cassette was excised from pUC18K by *Eco*RI-*Pst*I digestion and introduced in the *bcrD Eco*RI-*Sse*8387I sites. This construct generated an early stop in *bcrD* translation and allowed in-frame translation of the remaining 3' end of the mutated gene, avoiding possible polar effects on expression of downstream cistrons. The mutated *bcrD* gene with its flanking sequences, was excised by *Bgl*II-*Not*I cutting and subsequently inserted into the *Xba*I-*Eco*RI sites of the suicide plasmid pSS1129 (Stibitz, Methods Enzymol. (1994) 235:458-465), thanks to DNA adaptators. The resulting construction was named pAF214. pAF248 is a derivative of pAF214 that contained two additional unique *Spe*I and *Pac*I sites. These sites, included in a pair of complementary oligonucleotides, were introduced into the *Bam*HI site of pAF214. Other constructs included pAF245 and pAF246. PCR amplification of a 831 bp fragment covering the 5' region and the 4 first codons of *bcrD* was generated. This amplicon was further introduced into *Bam*HI-*Hin*DIII linearized pNM480 (Minton, Gene (1984) 31:269-273), in such a way that the *bcrD* initiation codon was placed in frame

with *lacZ*, used as a reporter gene. The resulting construct was named pAF245. Similarly, primers were designed for placing *lacZ* downstream of a 849 bp fragment that encompassed upstream *bscN* sequences including its 3 first codons. pAF246 was obtained by cloning this fragment in pNM480.

Transformations and allelic exchanges

B. pertussis cells, from a freshly saturated culture in 10 ml of SS medium, were washed and resuspended in 100µl of a cold 10% (v/v) glycerol solution. Up to 10 µg of supercoiled purified DNA in a maximum of 20 µl of water were added to 100 µl of the bacterial suspension. Cells and DNA were transferred to a prechilled 0.2 cm electroporation cuvette (Bio-Rad) and placed in a Gene Pulser apparatus (Bio-Rad). Pulses were achieved with settings of 25 µF, 2.5 kV, and 600 Ω, giving a time constant ranging from 11 to 14 ms.

After their initial isolation on BG plus gentamycin, pAF214 and pAF248 transformants that undergone a second recombination step were selected on streptomycin as described (Stibitz, *supra*). The null *bcrD* mutants were finally distinguished from revertants by their acquired resistance to kanamycin. The proper integration of the *aphA-3* was assessed by southern blot analysis. In contrast, introduction of pAF245 and pAF246 only required a single crossover selected on BG plus ampicillin. This recombination step led to the placement of the *lacZ* coding sequence under the control of the signals governing the transcription of *bcrD* and *bscN* respectively.

Mice model

After a two days growing on BG agar plates, wild type and mutant bacteria were recovered and resuspended in PBS at a concentration of 10^8 PFU ml^{-1}. 25 µl of the suspension were injected in each nostril of pentobarbital anaesthetized mice. Lungs colonization was assayed after 4 h, 3, 7, 14, 26, 39 and 45 days by treating both lungs of each mouse in an Ultraturax grinder and titrating the resuspended bacteria on BG agar plates.

β-galactosidase assay

0.5 ml of bacterial suspensions coming from liquid cultures grown to log phase (OD = 0.2), were assayed as described previously (Miller, (1972) "Experiments in molecular genetics." Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). We used the chromogenic substrate *o*-nitrophenyl-*β*-D-galactoside (ONPG) of Sigma.

Transcription of both bcrD *and* bscN *transcripts appear controlled by the* bvg *locus*

Most of the *Bordetella* virulence functions are controlled by the *bvg* locus. The Bvg^+ phase is characterized by the expression of virulence factors and is necessary for colonization of animal models. In contrast, the bacteria are avirulent in Bvg^- phase which can be induced by nicotinic acid or $MgSO_4$. We investigated the level of expression of two genes that belonged to distinct unit of transcription, i.e. *bcrD* and *bscN*, by using transcriptional fusions of *lacZ* into these genes. To this end, we isolated the mutants NIVh86 and NIVh87, which integrated pAF245 and pAF246 respectively. In the former mutant, a single recombination step led to the setting *lacZ* in place of the *bcrD* coding sequence, whereas in the latter, *lacZ* replaced *bscN*. The level of expression of both *bcrD* and *bscN* transcripts was assessed either in Bvg^+ or in Bvg^- phases. Both *B. pertussis* genes were weakly expressed *in vitro*. Additionally, however, these levels of expression appeared to be clearly modulated by the Bvg system. Indeed, whereas β-galactosidase could be assayed in Bvg^- conditions, no enzyme activity was detected in Bvg^- phase (table 5).

Table 5. β-galactosidase activity, in Miller units (Miller, *supra*), when *lacZ* is placed under the control of that direct the exrpession of *bcrD* or *bscN*.

phase transcript	Bvg^+	Bvg^-
bcrD	3.54	0.02
bscN	1.65	0.04

Example 4: Recombinant expression of effector protein vaccine candidates

In the discovered sequence, seven ORFs (*orf2* to *-8*) particularly fulfil certain criteria that make them good candidates as effector proteins and vaccine candidates. First, they appear surrounded by typical type III secretion (class I) genes, and therefore incontestably belong to the type III secretion locus. Furthermore, they don't display significant similarities with genes present in related type III systems from other organisms, and are therefore likely to be effector proteins specific for *Bordetella*. In addition to these ORFs *bopN*, *orf9* and *orf10* are also of particular interest as vaccine candidates. Despite the fact that these sequences do not fulfil the second criterium above (they have some similarity to *popN*, *pcrH* and *pcr4* of *Pseudomonas aeruginosa*), these products may also be exported by the specialized translocon. For these reasons, ten ORFs, i.e. *orf2* to *-10* and *bopN*, were selected for further analysis. To this end, ten pairs of primers (table 6) were designed for amplifying their corresponding ORF. The amplified ORFs were then cloned in the pCR-TOPO® T/A cloning system (Invitrogen) and their sequences were checked for errors putatively induced by the Taq DNA polymerase. Correct inserts were retrieved by *Eco*RI and *Bam*HI (or *Bgl*II - see table 6) cutting and transferred into the pMAL® vectors (New England Biolabs; Maina *et al.*, Gene (1988) 74:365-373), opened by *Eco*RI and *Bam*HI restriction. In these vectors, expression of the cloned inserts yields recombinant proteins fused to the maltose binding

protein (MBP) of *E. coli*. The MBP domain of the fusion protein provides a means for both detecting the expressed product and purifying it by affinity chromatography.

Four ORFs, namely *orf2*, *-4* and *-10* on the one hand, and *orf6* on the other, have been cloned into pMAL-c2E ® and pMAL-p2E® respectively. Transformed bacteria, grown in 300 ml of culture medium, were induced with IPTG (300 µM) and lysed in a French pressure cell. Insoluble material was pelleted by ultracentrifugation and discarded whereas the resulting supernatant was applied to an amylose resin. Fusion proteins that specifically bind to the amylose through their MBP domain, were further eluted by application of maltose 10 mM. This method allowed us to recover from 10 to 50 mg of each fusion protein (Fig. 6). The expressed *Bordetella* products may be separated from the MBP by utilising the enterokinase cleavage site between the *Bordetella* polypeptide and the MBP. The other ORFs should be expressable using a similar approach.

The secreted proteins will be analysed using standard techniques to confirm their functional and immunological properties. First, the immunogenicity of the secreted proteins will be assessed by investigating the presence of antibodies directed against these proteins in the serum of infected patients. In addition, their putative recognition as protective antigens will be based on challenge experiments, realized in a mouse model. Second, the biological properties of the effector proteins will be assessed by analysing their catalytic activities. For instance, it is expected that one of the secreted proteins would display a tyrosine phosphatase activity. Finally, the function of the effector proteins will be investigated by microinjecting the proteins into the cytoplasm of eukaryotic cells. This will allow us to display putative activities of inhibition of actin polymerisation, cytotoxicity or induction of apoptosis, i.e. those types of activities that have been assigned to effector proteins secreted by type III secretion systems discovered in other species.

Table 6. PCR primers used for amplifying the ORFs encoding vaccine candidates.

orf2	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** ATG CCG CAT ACC CTA CCC TCG
	complement	5'-TCT AGA GGA TCC GGC GAA TGG ATT TCT TGC TCG **TCA**
orf3	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** TCC AGC GCC GTA CCC GGC
	complement	5'-TCT AGA GGA TCC AGG GTA GGG TAT GCG GCA **TCA** TCC
orf4	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** AAT ACT GCC GAT AGG GCG CTG
	complement	5'-TCT AGA GGA TCC GGT ACG GCG CTG GAC ATG GCG **TC**
bopN	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** ACT CGT ATC GAT GCC GCC
	complement	5'-TCT AGA GGA TCC GCG CCC TAT CGG CAG TAT **TCA** TGC
orf5	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** GGG AGT CCT CGG AGA AGG AA
	complement	5'-TCT AGA GGA TCC ATA CTC CTT GTG CAG CGC **TTA** GCG
orf6	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** CAG GAG CAA GGC ATC CAA TC
	complement	5'-TCT AGA GGA TCC CAT GGA AGG CCT CCG CGC **TCA** GAC
orf7	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** TCT GTT TCT CCG ACT TCG CCC
	complement	5'-TCT AGA GGA TCC TGA AGG TTG GAG CCG GAC ACT **CAG**
orf8	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** ACC GTC ATG AGT ACG ACC ATA
	complement	5'-TCT AGA TCT TTC CTT GAG CGC CCG GCG **CTA** CA
orf9	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** ACT GTT CAT GAC GAC GCG
	complement	5'-TCT AGA GGA TCC GAG TCT GAG TGC ATG GAG **TTA** CTC C
orf10	direct	5'-GAG GAA TTC CAT ATG CCC ACC **ATG** CAC TCA GAC TCA GGT TCA GAT
	complement	5'-TCT AGA GGA TCC TCG CCG TCA GAT CCA AAT **TCA** TCC AG

Initiation and STOP codons of the corresponding ORF are written in bold. The cloning sites *Eco*RI, *Bam*HI or *Bgl*II are underlined. All but one of the complementary primers contain a *Bam*HI site. In the case of *orf8*, as it presents an internal *Bam*HI recognition sequence, a *Bgl*II site was preferred.

Bibliography

Allaoui, A., Woestyn, S., Sluiters, C., and Cornelis, G. R. (1994) YscU, A Yersinia enterocolitica inner membrane protein involved in Yop secretion. *J Bacteriol* **176**:4534-4542.

Bergman, T., Erickson, K., Galyov, E., Persson, C., and Wolf-Watz, H. (1994) The lcrB (yscN/U) gene cluster of Yersinia pseudotuberculosis is involved in Yop secretion and shows high homology to the spa gene clusters of Shigella flexneri and Salmonella typhimurium. *J Bacteriol* **176**:2619-2626.

Bogdanove, A. J., Wei, Z.-M., Zhao, L., and Beer, S. V. *Erwinia amylovora* secretes a Harpin via a type III pathway and contains a homolog of *yopN* of *Yersinia* spp. *J Bacteriol* **178**:1720-1730.

Fenselau, S., Balbo, I., and Bonas, U. (1992) Determinants of pathogenicity in Xanthomonas campestris pv. vesicatoria are related to proteins involved in secretion in bacterial pathogens of animals. *Mol Plant Interact* **5**:390-396.

Gyri (1995). *Mol. Microbiol.* **15**: 761-769.

Hueck, C. J. (1998) Type III protein secretion systems in Bacterial Pathogens of Animals and Plants. *Microb. Mol. Biol. Rev.* **62**: 379-433

Plano (1991). *J. Bact.* **173**: 7293-7303.

Ramakrishnan, G., Zhao, J. L., and Newton, A. (1991) The cell cycle-regulated flagellar gene flbF of Caulobacter crescentus is homologous to a virulence locus (lcrD) of Yersinia pestis. *J Bacteriol* **173**:7283-7292.

Yuk, M. H., Harvill, E. T., and Miller, J. F. (1998) The BvgAS virulence control system regulates type III secretion in *Bordetella bronchiseptica*. *Mol. Microbiol.* **28**:945-959.

SEQUENCE LISTING

<110> Alex Bollen
Alain Fauconnier
Edmond Godfroid

<120> Vaccine

<130> B45168

<160> 82

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2100
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(2100)

<400> 1

```
atg acg agc aag aaa tcc att cgc cgc ctg caa cgc gcg gtg gcg ctg       48
Met Thr Ser Lys Lys Ser Ile Arg Arg Leu Gln Arg Ala Val Ala Leu
 1               5                  10                  15

gcc acc agc cgc aac gac atc gta ctg gcc gtg ctc atc gtg gcg atc       96
Ala Thr Ser Arg Asn Asp Ile Val Leu Ala Val Leu Ile Val Ala Ile
            20                  25                  30

gtc ttc atg atg atc ctg ccg ttg ccc aca acg ctg gtc gac gtg ctg      144
Val Phe Met Met Ile Leu Pro Leu Pro Thr Thr Leu Val Asp Val Leu
        35                  40                  45

atc ggt gcg aac atg acg ctg tcg gca gtc ctg ctg atg gtc gcg atg      192
Ile Gly Ala Asn Met Thr Leu Ser Ala Val Leu Leu Met Val Ala Met
    50                  55                  60

tac ctg cct tcg ccc ctg gcg ttt tcc tcg ttc cct tcg gtc ctg ctg      240
Tyr Leu Pro Ser Pro Leu Ala Phe Ser Ser Phe Pro Ser Val Leu Leu
65                  70                  75                  80

gtc acc acg ctg ttc cgg ctg ggc atc tcc atc gcg acc acg cgg ctg      288
Val Thr Thr Leu Phe Arg Leu Gly Ile Ser Ile Ala Thr Thr Arg Leu
                85                  90                  95

atc ctg ctg caa ggc gat gcc ggc cac atc atc gag acc ttc ggc aac      336
Ile Leu Leu Gln Gly Asp Ala Gly His Ile Ile Glu Thr Phe Gly Asn
            100                 105                 110

ttc gtg gtg ggc ggc aac ctg atc gtc ggc ctg gtg gtt ttc ctc atc      384
Phe Val Val Gly Gly Asn Leu Ile Val Gly Leu Val Val Phe Leu Ile
        115                 120                 125

ctc acg atc gtg cag ttc gtg gtc atc acc aaa ggc gcg gag cgg gtg      432
Leu Thr Ile Val Gln Phe Val Val Ile Thr Lys Gly Ala Glu Arg Val
    130                 135                 140
```

```
gcc gaa gtc gcc gcg cgc ttc tcg ctg gac gcc atg ccc ggc aag cag      480
Ala Glu Val Ala Ala Arg Phe Ser Leu Asp Ala Met Pro Gly Lys Gln
145                 150                 155                 160

atg tcc atc gac gcg gac ttg cgc gcg ggc acc ata gac atg gac gaa      528
Met Ser Ile Asp Ala Asp Leu Arg Ala Gly Thr Ile Asp Met Asp Glu
                165                 170                 175

gcc cga cgc cga cgc cgt acg gtc gag aag gaa agc caa ctg tat ggc      576
Ala Arg Arg Arg Arg Arg Thr Val Glu Lys Glu Ser Gln Leu Tyr Gly
            180                 185                 190

gcc atg gac ggc gcg atg aag ttc gtc aag ggc gat gcc atc gcc ggc      624
Ala Met Asp Gly Ala Met Lys Phe Val Lys Gly Asp Ala Ile Ala Gly
        195                 200                 205

ctg atc atc gtt gcc gtc aac ctg ctt ggc ggc atg ctg gtc ggc gtg      672
Leu Ile Ile Val Ala Val Asn Leu Leu Gly Gly Met Leu Val Gly Val
    210                 215                 220

ctg cag cgc ggc ctg agc gcc ggc gag gcc gtg cag aca tat gcc atc      720
Leu Gln Arg Gly Leu Ser Ala Gly Glu Ala Val Gln Thr Tyr Ala Ile
225                 230                 235                 240

ctg acc ata ggc gac ggg ctc atc gcg cag atc ccg gcg ctg ttc atc      768
Leu Thr Ile Gly Asp Gly Leu Ile Ala Gln Ile Pro Ala Leu Phe Ile
                245                 250                 255

gcc atc tgc gcc gga atc atc gtg acg cgg gtg cag acc ggg gat ggc      816
Ala Ile Cys Ala Gly Ile Ile Val Thr Arg Val Gln Thr Gly Asp Gly
            260                 265                 270

ccc tcc aac gta ggc acc gac atc ggc gca caa gtg ctg gcg cag cct      864
Pro Ser Asn Val Gly Thr Asp Ile Gly Ala Gln Val Leu Ala Gln Pro
        275                 280                 285

cgc gcc ctg gtc att gcc ggc gcg atc tcg gca ggc ctg ggc ctc att      912
Arg Ala Leu Val Ile Ala Gly Ala Ile Ser Ala Gly Leu Gly Leu Ile
    290                 295                 300

ccc ggc atg ccc acg ctg gtc ttc ttc gcc ctg gcc gcc gcg gtg ggc      960
Pro Gly Met Pro Thr Leu Val Phe Phe Ala Leu Ala Ala Ala Val Gly
305                 310                 315                 320

acc atc ggt ttc gta ctg ctg cgc gca tcc cag cgt ccg ccc gaa ggc     1008
Thr Ile Gly Phe Val Leu Leu Arg Ala Ser Gln Arg Pro Pro Glu Gly
                325                 330                 335

gcc gag ccc gcg ctc gcc ggc atg gct gcc gac ggc cag ccc cgc acc     1056
Ala Glu Pro Ala Leu Ala Gly Met ala Ala Asp Gly Gln Pro Arg Thr
            340                 345                 350

cgc gcg ccg gcg gat ggg cag gcg gaa ttc gcc ccc acc gtc ccg ctg     1104
Arg Ala Pro Ala Asp Gly Gln Ala Glu Phe Ala Pro Thr Val Pro Leu
        355                 360                 365

atc atc gac gta gcc gcg cgg ctg cag ccc cgg ttc gag ccg gcc acc     1152
Ile Ile Asp Val Ala Ala Arg Leu Gln Pro Arg Phe Glu Pro Ala Thr
    370                 375                 380
```

```
ctc acc gac gat ctg ctg cag atc cgg cgg gcg ctc tat ttc gac ctg      1200
Leu Thr Asp Asp Leu Leu Gln Ile Arg Arg Ala Leu Tyr Phe Asp Leu
385                 390                 395                 400

ggc gtg ccg ttt ccc ggc atc cag ttg cgc ttc acc gaa gcg ctg gcc      1248
Gly Val Pro Phe Pro Gly Ile Gln Leu Arg Phe Thr Glu Ala Leu Ala
                405                 410                 415

gcc aat acc tac acc atc gtg ctg tcg gag atc ccg gtg gcg caa gga      1296
Ala Asn Thr Tyr Thr Ile Val Leu Ser Glu Ile Pro Val Ala Gln Gly
            420                 425                 430

atg ttg cgc gac gat gcc gtg ctg gtg cgg gac acc gag cag aac ctg      1344
Met Leu Arg Asp Asp Ala Val Leu Val Arg Asp Thr Glu Gln Asn Leu
        435                 440                 445

cag gcc ctg cgg atc gca tac gaa acg ggc gcg gcc ttt ctg ccc gat      1392
Gln Ala Leu Arg Ile Ala Tyr Glu Thr Gly Ala Ala Phe Leu Pro Asp
    450                 455                 460

acg ccc acg atc tgg gtt gcg gcc agt ctg acc ggc gcc ttg cgc gat      1440
Thr Pro Thr Ile Trp Val Ala Ala Ser Leu Thr Gly Ala Leu Arg Asp
465                 470                 475                 480

gca ggt att cct tac ctg ggt atc agc cag atc ctg act tgg cac ttg      1488
Ala Gly Ile Pro Tyr Leu Gly Ile Ser Gln Ile Leu Thr Trp His Leu
                485                 490                 495

gca tat gta ttg aaa aaa tat tca gcc gat ttc atc ggc atc cag gaa      1536
Ala Tyr Val Leu Lys Lys Tyr Ser Ala Asp Phe Ile Gly Ile Gln Glu
            500                 505                 510

acc cgg ttt ctg ctt tcg gcc atg gaa gaa cga ttt ccc gat ctg gtc      1584
Thr Arg Phe Leu Leu Ser Ala Met Glu Glu Arg Phe Pro Asp Leu Val
        515                 520                 525

aag gag tgc ctg cgc gtc atg ccg gtg cag aag att gcc gaa atc ctg      1632
Lys Glu Cys Leu Arg Val Met Pro Val Gln Lys Ile Ala Glu Ile Leu
    530                 535                 540

cag cgc ctt gtt tcc gaa gaa gtg tcg ata cgc aac ctg cgc gcc gtc      1680
Gln Arg Leu Val Ser Glu Glu Val Ser Ile Arg Asn Leu Arg Ala Val
545                 550                 555                 560

ctg gaa gcg ctg gtc gaa tgg ggc cag aag gaa aag gat acc gtc ctg      1728
Leu Glu Ala Leu Val Glu Trp Gly Gln Lys Glu Lys Asp Thr Val Leu
                565                 570                 575

ctt acg gag tat gtc cga atc gca ctc aag cgc tat atc agc cac aag      1776
Leu Thr Glu Tyr Val Arg Ile Ala Leu Lys Arg Tyr Ile Ser His Lys
            580                 585                 590

tac acc agc ggc cac aat atc ctg ccc gcc tac ctg ctg gcc ccc aag      1824
Tyr Thr Ser Gly His Asn Ile Leu Pro Ala Tyr Leu Leu Ala Pro Lys
        595                 600                 605

gtc gag gaa acc gtg cgc gcc gcc atc cgg cag acc gcc gcc ggc agt      1872
Val Glu Glu Thr Val Arg Ala Ala Ile Arg Gln Thr Ala Ala Gly Ser
    610                 615                 620
```

```
tat ctc gcc ctc gat ccg gac acg aca cgc cga ctg gtc gag cac atc      1920
Tyr Leu Ala Leu Asp Pro Asp Thr Thr Arg Arg Leu Val Glu His Ile
625                 630                 635                 640

cgt caa tgt gtc ggc gat ctg gcc gcc ggc gcg agc cgt ccc gtc ttg      1968
Arg Gln Cys Val Gly Asp Leu Ala Ala Gly Ala Ser Arg Pro Val Leu
                645                 650                 655

ctg acg tcg atg gac atc cgg cgc tac acg cgc aag atg ata gaa gcc      2016
Leu Thr Ser Met Asp Ile Arg Arg Tyr Thr Arg Lys Met Ile Glu Ala
            660                 665                 670

gat ctc tac gcc ctg ccg gtg ctg tcc tac cag gaa ctg acg ccg gag      2064
Asp Leu Tyr Ala Leu Pro Val Leu Ser Tyr Gln Glu Leu Thr Pro Glu
        675                 680                 685

atc aat gta cag ccc ctg ggc agg gtg gat cta tga                      2100
Ile Asn Val Gln Pro Leu Gly Arg Val Asp Leu  *
    690                 695
```

<210> 2
<211> 699
<212> PRT
<213> Bordetella pertussis

<400> 2

```
Met Thr Ser Lys Lys Ser Ile Arg Arg Leu Gln Arg Ala Val Ala Leu
 1               5                  10                  15
Ala Thr Ser Arg Asn Asp Ile Val Leu Ala Val Leu Ile Val Ala Ile
            20                  25                  30
Val Phe Met Met Ile Leu Pro Leu Pro Thr Thr Leu Val Asp Val Leu
        35                  40                  45
Ile Gly Ala Asn Met Thr Leu Ser Ala Val Leu Leu Met Val Ala Met
    50                  55                  60
Tyr Leu Pro Ser Pro Leu Ala Phe Ser Ser Phe Pro Ser Val Leu Leu
65                  70                  75                  80
Val Thr Thr Leu Phe Arg Leu Gly Ile Ser Ile Ala Thr Thr Arg Leu
                85                  90                  95
Ile Leu Leu Gln Gly Asp Ala Gly His Ile Ile Glu Thr Phe Gly Asn
            100                 105                 110
Phe Val Val Gly Gly Asn Leu Ile Val Gly Leu Val Val Phe Leu Ile
        115                 120                 125
Leu Thr Ile Val Gln Phe Val Val Ile Thr Lys Gly Ala Glu Arg Val
    130                 135                 140
Ala Glu Val Ala Ala Arg Phe Ser Leu Asp Ala Met Pro Gly Lys Gln
145                 150                 155                 160
Met Ser Ile Asp Ala Asp Leu Arg Ala Gly Thr Ile Asp Met Asp Glu
                165                 170                 175
Ala Arg Arg Arg Arg Arg Thr Val Glu Lys Glu Ser Gln Leu Tyr Gly
            180                 185                 190
Ala Met Asp Gly Ala Met Lys Phe Val Lys Gly Asp Ala Ile Ala Gly
        195                 200                 205
Leu Ile Ile Val Ala Val Asn Leu Leu Gly Gly Met Leu Val Gly Val
    210                 215                 220
Leu Gln Arg Gly Leu Ser Ala Gly Glu Ala Val Gln Thr Tyr Ala Ile
225                 230                 235                 240
Leu Thr Ile Gly Asp Gly Leu Ile Ala Gln Ile Pro Ala Leu Phe Ile
                245                 250                 255
```

```
Ala Ile Cys Ala Gly Ile Ile Val Thr Arg Val Gln Thr Gly Asp Gly
            260                 265                 270
Pro Ser Asn Val Gly Thr Asp Ile Gly Ala Gln Val Leu Ala Gln Pro
        275                 280                 285
Arg Ala Leu Val Ile Ala Gly Ala Ile Ser Ala Gly Leu Gly Leu Ile
    290                 295                 300
Pro Gly Met Pro Thr Leu Val Phe Phe Ala Leu Ala Ala Ala Val Gly
305                 310                 315                 320
Thr Ile Gly Phe Val Leu Leu Arg Ala Ser Gln Arg Pro Pro Glu Gly
                325                 330                 335
Ala Glu Pro Ala Leu Ala Gly Met ala Ala Asp Gly Gln Pro Arg Thr
            340                 345                 350
Arg Ala Pro Ala Asp Gly Gln Ala Glu Phe Ala Pro Thr Val Pro Leu
        355                 360                 365
Ile Ile Asp Val Ala Ala Arg Leu Gln Pro Arg Phe Glu Pro Ala Thr
    370                 375                 380
Leu Thr Asp Asp Leu Leu Gln Ile Arg Arg Ala Leu Tyr Phe Asp Leu
385                 390                 395                 400
Gly Val Pro Phe Pro Gly Ile Gln Leu Arg Phe Thr Glu Ala Leu Ala
                405                 410                 415
Ala Asn Thr Tyr Thr Ile Val Leu Ser Glu Ile Pro Val Ala Gln Gly
            420                 425                 430
Met Leu Arg Asp Asp Ala Val Leu Val Arg Asp Thr Glu Gln Asn Leu
        435                 440                 445
Gln Ala Leu Arg Ile Ala Tyr Glu Thr Gly Ala Ala Phe Leu Pro Asp
    450                 455                 460
Thr Pro Thr Ile Trp Val Ala Ala Ser Leu Thr Gly Ala Leu Arg Asp
465                 470                 475                 480
Ala Gly Ile Pro Tyr Leu Gly Ile Ser Gln Ile Leu Thr Trp His Leu
                485                 490                 495
Ala Tyr Val Leu Lys Lys Tyr Ser Ala Asp Phe Ile Gly Ile Gln Glu
            500                 505                 510
Thr Arg Phe Leu Leu Ser Ala Met Glu Glu Arg Phe Pro Asp Leu Val
        515                 520                 525
Lys Glu Cys Leu Arg Val Met Pro Val Gln Lys Ile Ala Glu Ile Leu
    530                 535                 540
Gln Arg Leu Val Ser Glu Glu Val Ser Ile Arg Asn Leu Arg Ala Val
545                 550                 555                 560
Leu Glu Ala Leu Val Glu Trp Gly Gln Lys Glu Lys Asp Thr Val Leu
                565                 570                 575
Leu Thr Glu Tyr Val Arg Ile Ala Leu Lys Arg Tyr Ile Ser His Lys
            580                 585                 590
Tyr Thr Ser Gly His Asn Ile Leu Pro Ala Tyr Leu Leu Ala Pro Lys
        595                 600                 605
Val Glu Glu Thr Val Arg Ala Ala Ile Arg Gln Thr Ala Ala Gly Ser
    610                 615                 620
Tyr Leu Ala Leu Asp Pro Asp Thr Thr Arg Arg Leu Val Glu His Ile
625                 630                 635                 640
Arg Gln Cys Val Gly Asp Leu Ala Ala Gly Ala Ser Arg Pro Val Leu
                645                 650                 655
Leu Thr Ser Met Asp Ile Arg Arg Tyr Thr Arg Lys Met Ile Glu Ala
            660                 665                 670
Asp Leu Tyr Ala Leu Pro Val Leu Ser Tyr Gln Glu Leu Thr Pro Glu
        675                 680                 685
Ile Asn Val Gln Pro Leu Gly Arg Val Asp Leu
    690                 695
```

<210> 3
<211> 486
<212> DNA

<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(486)

<400> 3

```
atg cca aag tca gcc gac cag ggc ggc tcc ccg gcg tca gct tcg cat        48
Met Pro Lys Ser Ala Asp Gln Gly Gly Ser Pro Ala Ser Ala Ser His
 1               5                  10                  15

gag gcg ttg cgc cat att ctc gac gca ggc gct tcg atg ggg ggc ttg        96
Glu Ala Leu Arg His Ile Leu Asp Ala Gly Ala Ser Met Gly Gly Leu
            20                  25                  30

cag ggg ttg gac gag gcg cag cag cag gcg ttg tac gcg atc ggt cat       144
Gln Gly Leu Asp Glu Ala Gln Gln Gln Ala Leu Tyr Ala Ile Gly His
        35                  40                  45

ggc gcc tac gaa cag ggg cgc tat gcc gac gcg ttg aaa atg ttc tgc       192
Gly Ala Tyr Glu Gln Gly Arg Tyr Ala Asp Ala Leu Lys Met Phe Cys
    50                  55                  60

ctg ctg gtc gcg tgc gat ccg ctg gaa gcc cgt tat ctg ctg gcc ctg       240
Leu Leu Val Ala Cys Asp Pro Leu Glu Ala Arg Tyr Leu Leu Ala Leu
65                  70                  75                  80

ggc gcc gcg gcc cag gag ctg ggg ctg tac gag cat gcc ttg cag caa       288
Gly Ala Ala Ala Gln Glu Leu Gly Leu Tyr Glu His Ala Leu Gln Gln
                85                  90                  95

tac gcg gcc gcg gcg gct ttg cag ttg gac tcc ccc agg ccc ctg ttg       336
Tyr Ala Ala Ala Ala Ala Leu Gln Leu Asp Ser Pro Arg Pro Leu Leu
            100                 105                 110

cat ggc gcc gag tgc ctg tat gcg ttg ggt cgt cgc cgc gac gcc ctg       384
His Gly Ala Glu Cys Leu Tyr Ala Leu Gly Arg Arg Arg Asp Ala Leu
        115                 120                 125

gat acg ctc gac atg gtg ctt gag ttg tgc ggc tcg ccg gag cgt gcg       432
Asp Thr Leu Asp Met Val Leu Glu Leu Cys Gly Ser Pro Glu Arg Ala
    130                 135                 140

gcc ctg cgc gaa cgg gcc gag ttg ctg cgc agg agc tat gca cgt gcc       480
Ala Leu Arg Glu Arg Ala Glu Leu Leu Arg Arg Ser Tyr Ala Arg Ala
145                 150                 155                 160

gac tga                                                              486
Asp  *
```

<210> 4
<211> 161
<212> PRT
<213> Bordetella pertussis

<400> 4

```
Met Pro Lys Ser Ala Asp Gln Gly Gly Ser Pro Ala Ser Ala Ser His
 1               5                  10                  15
```

Glu Ala Leu Arg His Ile Leu Asp Ala Gly Ala Ser Met Gly Gly Leu
20 25 30
Gln Gly Leu Asp Glu Ala Gln Gln Gln Ala Leu Tyr Ala Ile Gly His
35 40 45
Gly Ala Tyr Glu Gln Gly Arg Tyr Ala Asp Ala Leu Lys Met Phe Cys
50 55 60
Leu Leu Val Ala Cys Asp Pro Leu Glu Ala Arg Tyr Leu Leu Ala Leu
65 70 75 80
Gly Ala Ala Ala Gln Glu Leu Gly Leu Tyr Glu His Ala Leu Gln Gln
85 90 95
Tyr Ala Ala Ala Ala Ala Leu Gln Leu Asp Ser Pro Arg Pro Leu Leu
100 105 110
His Gly Ala Glu Cys Leu Tyr Ala Leu Gly Arg Arg Arg Asp Ala Leu
115 120 125
Asp Thr Leu Asp Met Val Leu Glu Leu Cys Gly Ser Pro Glu Arg Ala
130 135 140
Ala Leu Arg Glu Arg Ala Glu Leu Leu Arg Arg Ser Tyr Ala Arg Ala
145 150 155 160
Asp

<210> 5
<211> 1803
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1803)

<400> 5

```
atg gca ata ggt cgg ctt ggg tat ctt gtc cgc ggc gca tgg gcc ggg        48
Met ala Ile Gly Arg Leu Gly Tyr Leu Val Arg Gly Ala Trp Ala Gly
 1               5                  10                  15

ggt gtc atg ctg ttg gcg gcc ggt agc gcc tgg gcg gcg ccg aac tgg        96
Gly Val Met Leu Leu Ala Ala Gly Ser Ala Trp Ala Ala Pro Asn Trp
            20                  25                  30

cct ttg gcg ccg tat agc tac tac gcg cag cag cag agc ctg tcc gat       144
Pro Leu Ala Pro Tyr Ser Tyr Tyr Ala Gln Gln Gln Ser Leu Ser Asp
        35                  40                  45

gtg ctg cgc gag ttc gcc gca ggc ttc agc ctg gcg ttg caa cag ggc       192
Val Leu Arg Glu Phe Ala Ala Gly Phe Ser Leu Ala Leu Gln Gln Gly
    50                  55                  60

aaa ggg gtg caa ggc gtg gtc aat ggg cgt ttc aat gcg cgc aca ccc       240
Lys Gly Val Gln Gly Val Val Asn Gly Arg Phe Asn Ala Arg Thr Pro
65                  70                  75                  80

acg gag ttc atc gag cgt ctc agc ggc atc tat ggg ttc aac tgg ttc       288
Thr Glu Phe Ile Glu Arg Leu Ser Gly Ile Tyr Gly Phe Asn Trp Phe
                85                  90                  95

gtg cat gcc ggc acg ctg tat gtc agc cgc acc agc gac gtg gtt acc       336
Val His Ala Gly Thr Leu Tyr Val Ser Arg Thr Ser Asp Val Val Thr
            100                 105                 110

cgc gcg gtg gat gca gcc ggc gct tcg ccg tcg gcg ttg cgc cag gcc       384
```

```
Arg Ala Val Asp Ala Ala Gly Ala Ser Pro Ser Ala Leu Arg Gln Ala
        115                 120                 125

ttg ctg caa ctg ggc atc ctg gac gaa cgc ttc gga tgg gga gag ctg      432
Leu Leu Gln Leu Gly Ile Leu Asp Glu Arg Phe Gly Trp Gly Glu Leu
    130                 135                 140

ccg gcg caa ggc gtg gcc atg gtg tca ggg ccg ccg gcc tat gtc gcg      480
Pro Ala Gln Gly Val Ala Met Val Ser Gly Pro Pro Ala Tyr Val Ala
145                 150                 155                 160

ctg gtc gag cag gcg gta gcg gcg ttg ccc aag ggg gcc ggc aat cag      528
Leu Val Glu Gln Ala Val Ala Ala Leu Pro Lys Gly Ala Gly Asn Gln
                165                 170                 175

cag gtg gcg gtg ttt cgc ctc aag cat gct tcc gtg agc gac cgg gtg      576
Gln Val Ala Val Phe Arg Leu Lys His Ala Ser Val Ser Asp Arg Val
            180                 185                 190

atc cgt tat cga gac cag cag gta gtt acg ccg ggg atg gcc acc atg      624
Ile Arg Tyr Arg Asp Gln Gln Val Val Thr Pro Gly Met ala Thr Met
        195                 200                 205

ctg cgc caa ttg atc ctg ggg gcg ggg ccg ggc aac gac gcg gcg ctg      672
Leu Arg Gln Leu Ile Leu Gly Ala Gly Pro Gly Asn Asp Ala Ala Leu
    210                 215                 220

gcc gcg gtg gcg gcg ccg ctg cgg gaa aat ccg ccg gtg ttc ggc gat      720
Ala Ala Val Ala Ala Pro Leu Arg Glu Asn Pro Pro Val Phe Gly Asp
225                 230                 235                 240

gcg gca gct gac ggg aac gcg ccg ctc gct ggc gca gcc cag gca gcc      768
Ala Ala Ala Asp Gly Asn Ala Pro Leu Ala Gly Ala Ala Gln Ala Ala
                245                 250                 255

ggc cgg cgc ctg agc gag ccc agc gtg cag gcc gac acg cgc ctc aat      816
Gly Arg Arg Leu Ser Glu Pro Ser Val Gln Ala Asp Thr Arg Leu Asn
            260                 265                 270

gcc ttg atc gtg cag gat att ccc gaa cgg atg cca atc tac cgt gcc      864
Ala Leu Ile Val Gln Asp Ile Pro Glu Arg Met Pro Ile Tyr Arg Ala
        275                 280                 285

ctg atc gag cag ttg gat gtg ccc agc acc ctg atc gaa ata gag gcc      912
Leu Ile Glu Gln Leu Asp Val Pro Ser Thr Leu Ile Glu Ile Glu Ala
    290                 295                 300

atg atc gtg gac gtc aat acc gat ctg gtc aac gag ctg ggt gtc acc      960
Met Ile Val Asp Val Asn Thr Asp Leu Val Asn Glu Leu Gly Val Thr
305                 310                 315                 320

tgg ggg gcg cag atc gga acc acc agc ctg ggc tat ggc gat ctg ggg     1008
Trp Gly Ala Gln Ile Gly Thr Thr Ser Leu Gly Tyr Gly Asp Leu Gly
                325                 330                 335

ctg cgt ccc ggc aac ggc ctg ccc gtg gac ggc gcg gcg gcc gac ctg     1056
Leu Arg Pro Gly Asn Gly Leu Pro Val Asp Gly Ala Ala Ala Asp Leu
            340                 345                 350

gcg ccc gga acc ttg ggg atc agt gtc agt acc cgg ctg gcg gcg cgc     1104
```

```
Ala Pro Gly Thr Leu Gly Ile Ser Val Ser Thr Arg Leu Ala Ala Arg
          355                 360                 365

ttg cgt gcg ttg gag tcg gac ggg cag gcc aat atc ctg tct cag ccg      1152
Leu Arg Ala Leu Glu Ser Asp Gly Gln Ala Asn Ile Leu Ser Gln Pro
    370                 375                 380

tcc atc ctg acc gcc gac aac ctc ggc gcc atg ata gac ctg tcg gat      1200
Ser Ile Leu Thr Ala Asp Asn Leu Gly Ala Met Ile Asp Leu Ser Asp
385                 390                 395                 400

acc ttc tac att cgc acc ctg ggc gag cgc gta gcg aca gtc acg cct      1248
Thr Phe Tyr Ile Arg Thr Leu Gly Glu Arg Val Ala Thr Val Thr Pro
                405                 410                 415

gtc acg gtg ggt acg tcg ttg cgt gtg acg ccg cgc tat atc gcc gcc      1296
Val Thr Val Gly Thr Ser Leu Arg Val Thr Pro Arg Tyr Ile Ala Ala
            420                 425                 430

aag gga gga cgc cag gtg gaa ttg gcg atc gat atc gag gac gga cgg      1344
Lys Gly Gly Arg Gln Val Glu Leu Ala Ile Asp Ile Glu Asp Gly Arg
        435                 440                 445

gtc ttg cag gag tat ccc atc gat ggt ctg ccc cgg gtt cgg aaa agc      1392
Val Leu Gln Glu Tyr Pro Ile Asp Gly Leu Pro Arg Val Arg Lys Ser
    450                 455                 460

agc atc agc acg ctg gcg gtg gtg ggg gac gag cag acg ctg ctg atc      1440
Ser Ile Ser Thr Leu Ala Val Val Gly Asp Glu Gln Thr Leu Leu Ile
465                 470                 475                 480

ggc ggc tac aac aat cgc cgt gac gaa gag cag gtc gag aaa gtg ccg      1488
Gly Gly Tyr Asn Asn Arg Arg Asp Glu Glu Gln Val Glu Lys Val Pro
                485                 490                 495

ctg ctg gga gat atc ccc ggc ctg ggg ttc ttg ttc tcg agc aag tcc      1536
Leu Leu Gly Asp Ile Pro Gly Leu Gly Phe Leu Phe Ser Ser Lys Ser
            500                 505                 510

cgg gcg gta cag cgc cgc gag cgg ctg ttc ctg atc cgg ccg cgt gtc      1584
Arg Ala Val Gln Arg Arg Glu Arg Leu Phe Leu Ile Arg Pro Arg Val
        515                 520                 525

gtg gct atc gag ggc aag ccg gtc ttc agc ccc gtt gcg ggc acg tcg      1632
Val Ala Ile Glu Gly Lys Pro Val Phe Ser Pro Val Ala Gly Thr Ser
    530                 535                 540

cag gtg ttc atg agc acg ggt tgg ggc ggg cat ggc agc agc ctg agc      1680
Gln Val Phe Met Ser Thr Gly Trp Gly Gly His Gly Ser Ser Leu Ser
545                 550                 555                 560

att gca ccc ggc gag ggc ggg cat aca caa gtg cgt cat gat gcc cgg      1728
Ile Ala Pro Gly Glu Gly Gly His Thr Gln Val Arg His Asp Ala Arg
                565                 570                 575

gcg ggc agg ccg gtc cgg ctg gtg ccg gat tca ttg cat gtg gag tat      1776
Ala Gly Arg Pro Val Arg Leu Val Pro Asp Ser Leu His Val Glu Tyr
            580                 585                 590

ggc gag gcg ggg gag gcg tcg ccc tga                                  1803
```

Gly Glu Ala Gly Glu Ala Ser Pro *
595 600

<210> 6
<211> 600
<212> PRT
<213> Bordetella pertussis

<400> 6

Met	ala	Ile	Gly	Arg	Leu	Gly	Tyr	Leu	Val	Arg	Gly	Ala	Trp	Ala	Gly
1				5				10						15	
Gly	Val	Met	Leu	Leu	Ala	Ala	Gly	Ser	Ala	Trp	Ala	Ala	Pro	Asn	Trp
			20					25					30		
Pro	Leu	Ala	Pro	Tyr	Ser	Tyr	Tyr	Ala	Gln	Gln	Gln	Ser	Leu	Ser	Asp
		35					40					45			
Val	Leu	Arg	Glu	Phe	Ala	Ala	Gly	Phe	Ser	Leu	Ala	Leu	Gln	Gln	Gly
	50					55					60				
Lys	Gly	Val	Gln	Gly	Val	Val	Asn	Gly	Arg	Phe	Asn	Ala	Arg	Thr	Pro
65					70					75					80
Thr	Glu	Phe	Ile	Glu	Arg	Leu	Ser	Gly	Ile	Tyr	Gly	Phe	Asn	Trp	Phe
				85					90					95	
Val	His	Ala	Gly	Thr	Leu	Tyr	Val	Ser	Arg	Thr	Ser	Asp	Val	Val	Thr
			100					105					110		
Arg	Ala	Val	Asp	Ala	Ala	Gly	Ala	Ser	Pro	Ser	Ala	Leu	Arg	Gln	Ala
		115					120					125			
Leu	Leu	Gln	Leu	Gly	Ile	Leu	Asp	Glu	Arg	Phe	Gly	Trp	Gly	Glu	Leu
	130					135					140				
Pro	Ala	Gln	Gly	Val	Ala	Met	Val	Ser	Gly	Pro	Pro	Ala	Tyr	Val	Ala
145					150					155					160
Leu	Val	Glu	Gln	Ala	Val	Ala	Ala	Leu	Pro	Lys	Gly	Ala	Gly	Asn	Gln
				165					170					175	
Gln	Val	Ala	Val	Phe	Arg	Leu	Lys	His	Ala	Ser	Val	Ser	Asp	Arg	Val
			180					185					190		
Ile	Arg	Tyr	Arg	Asp	Gln	Gln	Val	Val	Thr	Pro	Gly	Met	ala	Thr	Met
		195					200					205			
Leu	Arg	Gln	Leu	Ile	Leu	Gly	Ala	Gly	Pro	Gly	Asn	Asp	Ala	Ala	Leu
	210					215					220				
Ala	Ala	Val	Ala	Ala	Pro	Leu	Arg	Glu	Asn	Pro	Pro	Val	Phe	Gly	Asp
225					230					235					240
Ala	Ala	Ala	Asp	Gly	Asn	Ala	Pro	Leu	Ala	Gly	Ala	Ala	Gln	Ala	Ala
				245					250					255	
Gly	Arg	Arg	Leu	Ser	Glu	Pro	Ser	Val	Gln	Ala	Asp	Thr	Arg	Leu	Asn
			260					265					270		
Ala	Leu	Ile	Val	Gln	Asp	Ile	Pro	Glu	Arg	Met	Pro	Ile	Tyr	Arg	Ala
		275					280					285			
Leu	Ile	Glu	Gln	Leu	Asp	Val	Pro	Ser	Thr	Leu	Ile	Glu	Ile	Glu	Ala
	290					295					300				
Met	Ile	Val	Asp	Val	Asn	Thr	Asp	Leu	Val	Asn	Glu	Leu	Gly	Val	Thr
305					310					315					320
Trp	Gly	Ala	Gln	Ile	Gly	Thr	Thr	Ser	Leu	Gly	Tyr	Gly	Asp	Leu	Gly
				325					330					335	
Leu	Arg	Pro	Gly	Asn	Gly	Leu	Pro	Val	Asp	Gly	Ala	Ala	Ala	Asp	Leu
			340					345					350		
Ala	Pro	Gly	Thr	Leu	Gly	Ile	Ser	Val	Ser	Thr	Arg	Leu	Ala	Ala	Arg
		355					360					365			
Leu	Arg	Ala	Leu	Glu	Ser	Asp	Gly	Gln	Ala	Asn	Ile	Leu	Ser	Gln	Pro
	370					375					380				
Ser	Ile	Leu	Thr	Ala	Asp	Asn	Leu	Gly	Ala	Met	Ile	Asp	Leu	Ser	Asp
385					390					395					400

```
Thr Phe Tyr Ile Arg Thr Leu Gly Glu Arg Val Ala Thr Val Thr Pro
                405                 410                 415
Val Thr Val Gly Thr Ser Leu Arg Val Thr Pro Arg Tyr Ile Ala Ala
            420                 425                 430
Lys Gly Gly Arg Gln Val Glu Leu Ala Ile Asp Ile Glu Asp Gly Arg
        435                 440                 445
Val Leu Gln Glu Tyr Pro Ile Asp Gly Leu Pro Arg Val Arg Lys Ser
    450                 455                 460
Ser Ile Ser Thr Leu Ala Val Val Gly Asp Glu Gln Thr Leu Leu Ile
465                 470                 475                 480
Gly Gly Tyr Asn Asn Arg Arg Asp Glu Glu Gln Val Glu Lys Val Pro
                485                 490                 495
Leu Leu Gly Asp Ile Pro Gly Leu Gly Phe Leu Phe Ser Ser Lys Ser
            500                 505                 510
Arg Ala Val Gln Arg Arg Glu Arg Leu Phe Leu Ile Arg Pro Arg Val
        515                 520                 525
Val Ala Ile Glu Gly Lys Pro Val Phe Ser Pro Val Ala Gly Thr Ser
    530                 535                 540
Gln Val Phe Met Ser Thr Gly Trp Gly Gly His Gly Ser Ser Leu Ser
545                 550                 555                 560
Ile Ala Pro Gly Glu Gly Gly His Thr Gln Val Arg His Asp Ala Arg
                565                 570                 575
Ala Gly Arg Pro Val Arg Leu Val Pro Asp Ser Leu His Val Glu Tyr
            580                 585                 590
Gly Glu Ala Gly Glu Ala Ser Pro
        595                 600
```

<210> 7
<211> 1281
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1281)

<400> 7

```
atg acg acg gcg ctg gaa ttc cgc gtg ctt tca ggc gca cag tgc atg        48
Met Thr Thr Ala Leu Glu Phe Arg Val Leu Ser Gly Ala Gln Cys Met
 1               5                  10                  15

gcg cgc tgc ccg gcc gtg cat ggc gcg cgc gtg ggc gcc aat ccg cat        96
Ala Arg Cys Pro Ala Val His Gly Ala Arg Val Gly Ala Asn Pro His
            20                  25                  30

tgc gat atc gtc ctg acc ggc gag gac atg ccc gaa gtg gcg gga tgg       144
Cys Asp Ile Val Leu Thr Gly Glu Asp Met Pro Glu Val Ala Gly Trp
        35                  40                  45

ctg gag atc gac cag tcc ggc tgg cgg ttg gcc ggc gcc gtg acg ccc       192
Leu Glu Ile Asp Gln Ser Gly Trp Arg Leu Ala Gly Ala Val Thr Pro
    50                  55                  60

ggc ctg gac gcc cag gcg ccg tgt ccg ccc gcg gcc ttc aac gaa ccc       240
Gly Leu Asp Ala Gln Ala Pro Cys Pro Pro Ala Ala Phe Asn Glu Pro
65                  70                  75                  80

gta gag ctg gga gcc gcc tgg atc acc gtg gcc gcc cct tcc gcg ccg       288
Val Glu Leu Gly Ala Ala Trp Ile Thr Val Ala Ala Pro Ser Ala Pro
                85                  90                  95
```

```
tgg ccc gcg ccg ccg gag ccg tgc ggc ccg gac ggc agc gac aca gcc      336
Trp Pro Ala Pro Pro Glu Pro Cys Gly Pro Asp Gly Ser Asp Thr Ala
            100                 105                 110

ttg cac gac gtc cct ggc tcg aca agc ccg ccg tcc gtc gct gcc ctc      384
Leu His Asp Val Pro Gly Ser Thr Ser Pro Pro Ser Val Ala Ala Leu
        115                 120                 125

atg ccg cgc cga cgt gca gga cgg ccc tgg ctg gcg ctg ggc gcg gcc      432
Met Pro Arg Arg Arg Ala Gly Arg Pro Trp Leu Ala Leu Gly Ala Ala
    130                 135                 140

gcg gcc gtc ctg ctg gtc ggc ctg gcc acg gcg ctg gtt tcc gtg acc      480
Ala Ala Val Leu Leu Val Gly Leu Ala Thr Ala Leu Val Ser Val Thr
145                 150                 155                 160

aca ccc gcc acg ccg ccg gcc gcg ccg ccc cca acg ccc acc gcg ccg      528
Thr Pro Ala Thr Pro Pro Ala Ala Pro Pro Pro Thr Pro Thr Ala Pro
                165                 170                 175

ctg gtc cgc gcc gcg gcg ctc atc gac agc ctg ggc ctt acc gag caa      576
Leu Val Arg Ala Ala Ala Leu Ile Asp Ser Leu Gly Leu Thr Glu Gln
            180                 185                 190

tta caa gcg gcc tac ggc cgt ggc ggc gtg ctc acc gtg acc gga tgg      624
Leu Gln Ala Ala Tyr Gly Arg Gly Gly Val Leu Thr Val Thr Gly Trp
        195                 200                 205

gtg cac gac gag acg gaa ttc gct cgg gtc gcc agg gcg ttg gcg caa      672
Val His Asp Glu Thr Glu Phe Ala Arg Val Ala Arg Ala Leu Ala Gln
    210                 215                 220

ctt gcg cca cgg cct gcc atg cag gta agc agg cag gac gag gcc agg      720
Leu Ala Pro Arg Pro Ala Met Gln Val Ser Arg Gln Asp Glu Ala Arg
225                 230                 235                 240

gcc ctg gcc tgc gat gtc ctg gcg aca ttc ggg gtg cgc tac atg gcg      768
Ala Leu Ala Cys Asp Val Leu Ala Thr Phe Gly Val Arg Tyr Met ala
                245                 250                 255

cgc ccg tac ggc aat ggc cgc ctg gcg atc tcg ggc atc gcc agc gat      816
Arg Pro Tyr Gly Asn Gly Arg Leu Ala Ile Ser Gly Ile Ala Ser Asp
            260                 265                 270

gcg cac gaa cgc gcc gcg gcg ctg cat gcg gtg cgc atg cgc ctg ccg      864
Ala His Glu Arg Ala Ala Ala Leu His Ala Val Arg Met Arg Leu Pro
        275                 280                 285

ggc atg acg atc ctc ggt cgc gat gta cgc ctg gcc gac gag gtc tcg      912
Gly Met Thr Ile Leu Gly Arg Asp Val Arg Leu Ala Asp Glu Val Ser
    290                 295                 300

gcc cag ttc gcg gcc cag ctg gcc gac gaa cgc ctc gac ggc gtc aag      960
Ala Gln Phe Ala Ala Gln Leu Ala Asp Glu Arg Leu Asp Gly Val Lys
305                 310                 315                 320

ctc agc tgg cac gcc gac cgc ctg gac gca gat ccc ggc gga ttg gcg     1008
Leu Ser Trp His Ala Asp Arg Leu Asp Ala Asp Pro Gly Gly Leu Ala
                325                 330                 335
```

```
gca ggc cgc atg gcg cgc ctg cgc gag ctg gtg gcc gcg ttc aac cag      1056
Ala Gly Arg Met ala Arg Leu Arg Glu Leu Val Ala Ala Phe Asn Gln
            340                 345                 350

cgc aac tac gac gtc gtc cgg ctg ccg gcc acc gcc gcg cgc gcg acg      1104
Arg Asn Tyr Asp Val Val Arg Leu Pro Ala Thr Ala Ala Arg Ala Thr
        355                 360                 365

cgg gat cac gtg ccg ttc gag ata cgc agt gtc gtg agc ggc ccg caa      1152
Arg Asp His Val Pro Phe Glu Ile Arg Ser Val Val Ser Gly Pro Gln
    370                 375                 380

ccg tac ctg atg ctg gcc gat ggc agc cgc ctc ctg gtg ggc gga ctg      1200
Pro Tyr Leu Met Leu Ala Asp Gly Ser Arg Leu Leu Val Gly Gly Leu
385                 390                 395                 400

cgg gac cag tat cgc ctt acc gcc atc gaa tcc ggc cgc ctg gtc ttc      1248
Arg Asp Gln Tyr Arg Leu Thr Ala Ile Glu Ser Gly Arg Leu Val Phe
                405                 410                 415

gat ggt ccc gaa ccg gtc atc gtg acg cga tga                          1281
Asp Gly Pro Glu Pro Val Ile Val Thr Arg  *
            420                 425
```

<210> 8
<211> 426
<212> PRT
<213> Bordetella pertussis

<400> 8

```
Met Thr Thr Ala Leu Glu Phe Arg Val Leu Ser Gly Ala Gln Cys Met
 1               5                  10                  15
Ala Arg Cys Pro Ala Val His Gly Ala Arg Val Gly Ala Asn Pro His
            20                  25                  30
Cys Asp Ile Val Leu Thr Gly Glu Asp Met Pro Glu Val Ala Gly Trp
        35                  40                  45
Leu Glu Ile Asp Gln Ser Gly Trp Arg Leu Ala Gly Ala Val Thr Pro
    50                  55                  60
Gly Leu Asp Ala Gln Ala Pro Cys Pro Pro Ala Ala Phe Asn Glu Pro
65                  70                  75                  80
Val Glu Leu Gly Ala Ala Trp Ile Thr Val Ala Ala Pro Ser Ala Pro
                85                  90                  95
Trp Pro Ala Pro Pro Glu Pro Cys Gly Pro Asp Gly Ser Asp Thr Ala
            100                 105                 110
Leu His Asp Val Pro Gly Ser Thr Ser Pro Pro Ser Val Ala Ala Leu
        115                 120                 125
Met Pro Arg Arg Arg Ala Gly Arg Pro Trp Leu Ala Leu Gly Ala Ala
    130                 135                 140
Ala Ala Val Leu Leu Val Gly Leu Ala Thr Ala Leu Val Ser Val Thr
145                 150                 155                 160
Thr Pro Ala Thr Pro Pro Ala Ala Pro Pro Pro Thr Pro Thr Ala Pro
                165                 170                 175
Leu Val Arg Ala Ala Ala Leu Ile Asp Ser Leu Gly Leu Thr Glu Gln
            180                 185                 190
Leu Gln Ala Ala Tyr Gly Arg Gly Gly Val Leu Thr Val Thr Gly Trp
        195                 200                 205
Val His Asp Glu Thr Glu Phe Ala Arg Val Ala Arg Ala Leu Ala Gln
    210                 215                 220
```

```
Leu Ala Pro Arg Pro Ala Met Gln Val Ser Arg Gln Asp Glu Ala Arg
225                 230                 235                 240
Ala Leu Ala Cys Asp Val Leu Ala Thr Phe Gly Val Arg Tyr Met ala
                245                 250                 255
Arg Pro Tyr Gly Asn Gly Arg Leu Ala Ile Ser Gly Ile Ala Ser Asp
            260                 265                 270
Ala His Glu Arg Ala Ala Ala Leu His Ala Val Arg Met Arg Leu Pro
        275                 280                 285
Gly Met Thr Ile Leu Gly Arg Asp Val Arg Leu Ala Asp Glu Val Ser
    290                 295                 300
Ala Gln Phe Ala Ala Gln Leu Ala Asp Glu Arg Leu Asp Gly Val Lys
305                 310                 315                 320
Leu Ser Trp His Ala Asp Arg Leu Asp Ala Asp Pro Gly Gly Leu Ala
                325                 330                 335
Ala Gly Arg Met ala Arg Leu Arg Glu Leu Val Ala Ala Phe Asn Gln
            340                 345                 350
Arg Asn Tyr Asp Val Val Arg Leu Pro Ala Thr Ala Ala Arg Ala Thr
        355                 360                 365
Arg Asp His Val Pro Phe Glu Ile Arg Ser Val Val Ser Gly Pro Gln
    370                 375                 380
Pro Tyr Leu Met Leu Ala Asp Gly Ser Arg Leu Leu Val Gly Gly Leu
385                 390                 395                 400
Arg Asp Gln Tyr Arg Leu Thr Ala Ile Glu Ser Gly Arg Leu Val Phe
                405                 410                 415
Asp Gly Pro Glu Pro Val Ile Val Thr Arg
            420                 425
```

<210> 9
<211> 300
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(300)

<400> 9

```
atg agt aca tct gtt cta gcc ctg acc gaa ctg gaa gtg cgc ctg gca        48
Met Ser Thr Ser Val Leu Ala Leu Thr Glu Leu Glu Val Arg Leu Ala
 1               5                  10                  15

tcg ccg ggc ggt tcc gcc ttg cgc gac acc ttg ctg tcg cag ctt ggc        96
Ser Pro Gly Gly Ser Ala Leu Arg Asp Thr Leu Leu Ser Gln Leu Gly
            20                  25                  30

gaa ctg gag aca cgt ttg cgc gtc cgc ctg cac gat ggc gtg ggg cgc       144
Glu Leu Glu Thr Arg Leu Arg Val Arg Leu His Asp Gly Val Gly Arg
        35                  40                  45

gac acc tat ccc gta tgg cgc gac gcg ctg gcg gcc gcc acc gcg gcc       192
Asp Thr Tyr Pro Val Trp Arg Asp Ala Leu Ala Ala Ala Thr Ala Ala
    50                  55                  60

cgg cag gta ttg ctg cag cgc ccg acc ggg ccg gac aac cct ccg gcg       240
Arg Gln Val Leu Leu Gln Arg Pro Thr Gly Pro Asp Asn Pro Pro Ala
65                  70                  75                  80

tca gtc ttg acg cgc ctg agc aat gaa caa tgc gcc gaa gga gac aag       288
Ser Val Leu Thr Arg Leu Ser Asn Glu Gln Cys Ala Glu Gly Asp Lys
                85                  90                  95
```

```
cat ggc cat taa                                                          300
His Gly His  *
```

<210> 10
<211> 99
<212> PRT
<213> Bordetella pertussis

<400> 10

```
Met Ser Thr Ser Val Leu Ala Leu Thr Glu Leu Glu Val Arg Leu Ala
 1               5                  10                  15
Ser Pro Gly Gly Ser Ala Leu Arg Asp Thr Leu Leu Ser Gln Leu Gly
            20                  25                  30
Glu Leu Glu Thr Arg Leu Arg Val Arg Leu His Asp Gly Val Gly Arg
        35                  40                  45
Asp Thr Tyr Pro Val Trp Arg Asp Ala Leu Ala Ala Ala Thr Ala Ala
    50                  55                  60
Arg Gln Val Leu Leu Gln Arg Pro Thr Gly Pro Asp Asn Pro Pro Ala
65                  70                  75                  80
Ser Val Leu Thr Arg Leu Ser Asn Glu Gln Cys Ala Glu Gly Asp Lys
                85                  90                  95
His Gly His
```

<210> 11
<211> 267
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(267)

<400> 11

```
atg gcc att aac ctg gga ggc gac gca ggc cga gtg acc atg cag agc      48
Met ala Ile Asn Leu Gly Gly Asp Ala Gly Arg Val Thr Met Gln Ser
 1               5                  10                  15

gtc aac cag gcg gtc aat acg cgg ctg aac gct cac gaa cgc gac ctg      96
Val Asn Gln Ala Val Asn Thr Arg Leu Asn Ala His Glu Arg Asp Leu
            20                  25                  30

cgc agc cgc ctg gag gcg ctc agc gcg cgc gga gac ggc gcg gtc agc     144
Arg Ser Arg Leu Glu Ala Leu Ser Ala Arg Gly Asp Gly Ala Val Ser
        35                  40                  45

acg tcc gac ctg ctg atc gtg caa cag gaa atg caa tcg tgg gtc gtg     192
Thr Ser Asp Leu Leu Ile Val Gln Gln Glu Met Gln Ser Trp Val Val
    50                  55                  60

atg atc gat cta cag agc acg gtg gtc aag cag gtc gcg gat tcg ctc     240
Met Ile Asp Leu Gln Ser Thr Val Val Lys Gln Val Ala Asp Ser Leu
65                  70                  75                  80

aag ggc gtc ata cag aag gcg agt tga                                 267
Lys Gly Val Ile Gln Lys Ala Ser  *
                85
```

<210> 12
<211> 88
<212> PRT
<213> Bordetella pertussis

<400> 12

```
Met ala Ile Asn Leu Gly Gly Asp Ala Gly Arg Val Thr Met Gln Ser
 1               5                  10                  15
Val Asn Gln Ala Val Asn Thr Arg Leu Asn Ala His Glu Arg Asp Leu
            20                  25                  30
Arg Ser Arg Leu Glu Ala Leu Ser Ala Arg Gly Asp Gly Ala Val Ser
        35                  40                  45
Thr Ser Asp Leu Leu Ile Val Gln Gln Glu Met Gln Ser Trp Val Val
    50                  55                  60
Met Ile Asp Leu Gln Ser Thr Val Val Lys Gln Val Ala Asp Ser Leu
65                  70                  75                  80
Lys Gly Val Ile Gln Lys Ala Ser
                85
```

<210> 13
<211> 327
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(327)

<400> 13

```
atg gcg gat cag gcg cgc ttt gaa ttg gcg ctg ggc gag atg ccc ggc      48
Met ala Asp Gln Ala Arg Phe Glu Leu Ala Leu Gly Glu Met Pro Gly
 1               5                  10                  15

gca tcg gcc ccg aac ggg gcg atc gcc ctg gcg ccg gtc gcg ctc gac      96
Ala Ser Ala Pro Asn Gly Ala Ile Ala Leu Ala Pro Val Ala Leu Asp
            20                  25                  30

gag ccg ctg ggc cgt cgc att ctt gga cag ttg cgc ggc ggc ctg gcc     144
Glu Pro Leu Gly Arg Arg Ile Leu Gly Gln Leu Arg Gly Gly Leu Ala
        35                  40                  45

gat gtg gca gga aaa tgg cgg gcg gtg cag acg ggc ttg gcc gag gtg     192
Asp Val Ala Gly Lys Trp Arg Ala Val Gln Thr Gly Leu Ala Glu Val
    50                  55                  60

agc cag gcg cct acc gtg gtg ggt atg ctc gat ctg cag gcc agg ttg     240
Ser Gln Ala Pro Thr Val Val Gly Met Leu Asp Leu Gln Ala Arg Leu
65                  70                  75                  80

cta cag gca tcc gtg gag tac gag ttg gtg ggc aag gca ata ggg cgc     288
Leu Gln Ala Ser Val Glu Tyr Glu Leu Val Gly Lys Ala Ile Gly Arg
                85                  90                  95

gcc acc caa aac gtc gat acg ctg gcg aga atg tca tga                 327
Ala Thr Gln Asn Val Asp Thr Leu Ala Arg Met Ser  *
            100                 105
```

<210> 14
<211> 108
<212> PRT
<213> Bordetella pertussis

<400> 14

```
Met ala Asp Gln Ala Arg Phe Glu Leu Ala Leu Gly Glu Met Pro Gly
 1               5                  10                  15
Ala Ser Ala Pro Asn Gly Ala Ile Ala Leu Ala Pro Val Ala Leu Asp
            20                  25                  30
Glu Pro Leu Gly Arg Arg Ile Leu Gly Gln Leu Arg Gly Gly Leu Ala
        35                  40                  45
Asp Val Ala Gly Lys Trp Arg Ala Val Gln Thr Gly Leu Ala Glu Val
    50                  55                  60
Ser Gln Ala Pro Thr Val Val Gly Met Leu Asp Leu Gln Ala Arg Leu
65                  70                  75                  80
Leu Gln Ala Ser Val Glu Tyr Glu Leu Val Gly Lys Ala Ile Gly Arg
                85                  90                  95
Ala Thr Gln Asn Val Asp Thr Leu Ala Arg Met Ser
            100                 105
```

<210> 15
<211> 825
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(825)

<400> 15

```
atg aac gcc atc ggg gcg atc caa cgg tat cgg cgc ggc gcg gga tgg      48
Met Asn Ala Ile Gly Ala Ile Gln Arg Tyr Arg Arg Gly Ala Gly Trp
 1               5                  10                  15

gcg gcc ctg gtg ctc gcc ctg gcg ctg ctg gcc ggc tgc ggt gcc cgc      96
Ala Ala Leu Val Leu Ala Leu Ala Leu Leu Ala Gly Cys Gly Ala Arg
            20                  25                  30

gtc gag ctg ttg ggc gcg gcg ccc gag aac gaa gcc aac gaa gta ttg     144
Val Glu Leu Leu Gly Ala Ala Pro Glu Asn Glu Ala Asn Glu Val Leu
        35                  40                  45

gcg gcg ctg ctc gag gca ggc atc gct gcg cag aag cag tcc ggc aag     192
Ala Ala Leu Leu Glu Ala Gly Ile Ala Ala Gln Lys Gln Ser Gly Lys
    50                  55                  60

gcc ggc tac gcg gtt tcg gtg ccg gcc gag gcg gtg gcc cgg tcg ctg     240
Ala Gly Tyr Ala Val Ser Val Pro Ala Glu Ala Val Ala Arg Ser Leu
65                  70                  75                  80

gag atc ctg cgc gca agc ggc ctg ccc cgc gag cag ttc gac gga atg     288
Glu Ile Leu Arg Ala Ser Gly Leu Pro Arg Glu Gln Phe Asp Gly Met
                85                  90                  95

gga cgc ata ttc cgc aag gaa ggc ctg gtt tca tct ccg ctc gaa gag     336
Gly Arg Ile Phe Arg Lys Glu Gly Leu Val Ser Ser Pro Leu Glu Glu
            100                 105                 110

cgc gcc cgc tac att tat gcg ctg tct cag gaa ttg gcc gac acc ctg     384
```

```
Arg Ala Arg Tyr Ile Tyr Ala Leu Ser Gln Glu Leu Ala Asp Thr Leu
        115                 120                 125

tcg cag atc gac ggc gtg ctc agc gcc cgc gtg cac gtg gtg ctt ccc      432
Ser Gln Ile Asp Gly Val Leu Ser Ala Arg Val His Val Val Leu Pro
    130                 135                 140

gaa cgc ggc gcg gtc ggc gag ccg gcc acc cct tcg acg gca ggg gtg      480
Glu Arg Gly Ala Val Gly Glu Pro Ala Thr Pro Ser Thr Ala Gly Val
145                 150                 155                 160

ttt ctc aag tac cgc gac gga cag agc ctc gac gcg ctc gtg ccc gag      528
Phe Leu Lys Tyr Arg Asp Gly Gln Ser Leu Asp Ala Leu Val Pro Glu
                165                 170                 175

atc cgc aag ctg gtc acg cat gcc atc ccg ggc ctg gcc gag gac cgt      576
Ile Arg Lys Leu Val Thr His Ala Ile Pro Gly Leu Ala Glu Asp Arg
            180                 185                 190

gta tcg gtt gcc ctg gtg gtg gcc cag ccc gtt cag gcc gca ccc gcg      624
Val Ser Val Ala Leu Val Val Ala Gln Pro Val Gln Ala Ala Pro Ala
        195                 200                 205

ccg gtc gcg tgg cgc cgc gtg ctt ggc gta cag gtc gcg gac gga tcg      672
Pro Val Ala Trp Arg Arg Val Leu Gly Val Gln Val Ala Asp Gly Ser
    210                 215                 220

gtc ctg aga ttt tcg ctg ttg ctg ctg ttg ttg ccg gtg ctg tgc ctg      720
Val Leu Arg Phe Ser Leu Leu Leu Leu Leu Leu Pro Val Leu Cys Leu
225                 230                 235                 240

ata gtg gcg ggg gcc gcg ctc tac gtc tgg cgc acg cgc tgg tcc cgc      768
Ile Val Ala Gly Ala Ala Leu Tyr Val Trp Arg Thr Arg Trp Ser Arg
                245                 250                 255

ggc gaa ggg cgc ggc ggc gct ggc gcc ggc gcc acg gaa gga gcc ggg      816
Gly Glu Gly Arg Gly Gly Ala Gly Ala Gly Ala Thr Glu Gly Ala Gly
            260                 265                 270

cat gac tga                                                         825
His Asp  *
```

<210> 16
<211> 274
<212> PRT
<213> Bordetella pertussis

<400> 16

```
Met Asn Ala Ile Gly Ala Ile Gln Arg Tyr Arg Arg Gly Ala Gly Trp
 1               5                  10                  15
Ala Ala Leu Val Leu Ala Leu Ala Leu Leu Ala Gly Cys Gly Ala Arg
            20                  25                  30
Val Glu Leu Leu Gly Ala Ala Pro Glu Asn Glu Ala Asn Glu Val Leu
        35                  40                  45
Ala Ala Leu Leu Glu Ala Gly Ile Ala Ala Gln Lys Gln Ser Gly Lys
    50                  55                  60
Ala Gly Tyr Ala Val Ser Val Pro Ala Glu Ala Val Ala Arg Ser Leu
65                  70                  75                  80
```

```
Glu Ile Leu Arg Ala Ser Gly Leu Pro Arg Glu Gln Phe Asp Gly Met
                85                  90                  95
Gly Arg Ile Phe Arg Lys Glu Gly Leu Val Ser Ser Pro Leu Glu Glu
            100                 105                 110
Arg Ala Arg Tyr Ile Tyr Ala Leu Ser Gln Glu Leu Ala Asp Thr Leu
        115                 120                 125
Ser Gln Ile Asp Gly Val Leu Ser Ala Arg Val His Val Val Leu Pro
    130                 135                 140
Glu Arg Gly Ala Val Gly Glu Pro Ala Thr Pro Ser Thr Ala Gly Val
145                 150                 155                 160
Phe Leu Lys Tyr Arg Asp Gly Gln Ser Leu Asp Ala Leu Val Pro Glu
                165                 170                 175
Ile Arg Lys Leu Val Thr His Ala Ile Pro Gly Leu Ala Glu Asp Arg
            180                 185                 190
Val Ser Val Ala Leu Val Val Ala Gln Pro Val Gln Ala Ala Pro Ala
        195                 200                 205
Pro Val Ala Trp Arg Arg Val Leu Gly Val Gln Val Ala Asp Gly Ser
    210                 215                 220
Val Leu Arg Phe Ser Leu Leu Leu Leu Leu Leu Pro Val Leu Cys Leu
225                 230                 235                 240
Ile Val Ala Gly Ala Ala Leu Tyr Val Trp Arg Thr Arg Trp Ser Arg
                245                 250                 255
Gly Glu Gly Arg Gly Gly Ala Gly Ala Gly Ala Thr Glu Gly Ala Gly
            260                 265                 270
His Asp
```

<210> 17
<211> 663
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(663)

<400> 17

```
atg act gag gcg agc gtg ctg ctt tcc gag cgg ctc atg ata ttc aat   48
Met Thr Glu Ala Ser Val Leu Leu Ser Glu Arg Leu Met Ile Phe Asn
 1               5                  10                  15

ctc ctg ccc agc ctg acc ctg cat gcc agt cgc cac gac gag atg ttt   96
Leu Leu Pro Ser Leu Thr Leu His Ala Ser Arg His Asp Glu Met Phe
            20                  25                  30

cca gcc gat tgg gtg cgc gcg ttg tgc aat gcc gac gcg gcg ttg gcc  144
Pro Ala Asp Trp Val Arg Ala Leu Cys Asn Ala Asp Ala Ala Leu Ala
        35                  40                  45

aac gcg tgg cat cgc cat tgg tcg cgc tgg atc ttg tgc gaa ttg ggc  192
Asn Ala Trp His Arg His Trp Ser Arg Trp Ile Leu Cys Glu Leu Gly
    50                  55                  60

ctg ctg aac cag ccg gtc ctg agc ctc gat ccg ccg cag ttg aag gtc  240
Leu Leu Asn Gln Pro Val Leu Ser Leu Asp Pro Pro Gln Leu Lys Val
65                  70                  75                  80

gcg cta ttg tcc acg gac gcc ttg cgg acc tgc gcc gcc cat gcg gga  288
Ala Leu Leu Ser Thr Asp Ala Leu Arg Thr Cys Ala Ala His Ala Gly
                85                  90                  95
```

```
gcg ctg ctg tgc gcg ccg cgc ctg cga cgc gcg ata gac ggc gcc gag      336
Ala Leu Leu Cys Ala Pro Arg Leu Arg Arg Ala Ile Asp Gly Ala Glu
            100                 105                 110

gtc cgt acc ttg cat gcc gcg ctc ggg cgc gat gtg atg aat ttc gcc      384
Val Arg Thr Leu His Ala Ala Leu Gly Arg Asp Val Met Asn Phe Ala
        115                 120                 125

gtg tct tcc gcg gcg cgg gcc ctg cat gac ggg ctc gcc gcc agt tcg      432
Val Ser Ser Ala Ala Arg Ala Leu His Asp Gly Leu Ala Ala Ser Ser
    130                 135                 140

gac tgg acc ctg gcc gcc acg gtc cag gcg gcg cag aaa ctg ggc tgg      480
Asp Trp Thr Leu Ala Ala Thr Val Gln Ala Ala Gln Lys Leu Gly Trp
145                 150                 155                 160

gcc ctg ctg cgc gac gcc gtg cag ggc gcc gcc gac gag ata gcg ctg      528
Ala Leu Leu Arg Asp Ala Val Gln Gly Ala Ala Asp Glu Ile Ala Leu
                165                 170                 175

cgt tgc gcg ctg aag ttg ccg cgc gac ctt gat ccc gcg ccc gtc ctg      576
Arg Cys Ala Leu Lys Leu Pro Arg Asp Leu Asp Pro Ala Pro Val Leu
            180                 185                 190

ccg ccc gag gcg gcg ctt gcg ctg gtg ctg tcc atg ctc gaa atc ctg      624
Pro Pro Glu Ala Ala Leu Ala Leu Val Leu Ser Met Leu Glu Ile Leu
        195                 200                 205

gat gca gaa tgg ctt tcc tcg ttc ccc gcc caa gcc tga                  663
Asp Ala Glu Trp Leu Ser Ser Phe Pro Ala Gln Ala  *
    210                 215                 220
```

<210> 18
<211> 220
<212> PRT
<213> Bordetella pertussis

<400> 18

```
Met Thr Glu Ala Ser Val Leu Leu Ser Glu Arg Leu Met Ile Phe Asn
 1               5                  10                  15
Leu Leu Pro Ser Leu Thr Leu His Ala Ser Arg His Asp Glu Met Phe
            20                  25                  30
Pro Ala Asp Trp Val Arg Ala Leu Cys Asn Ala Asp Ala Ala Leu Ala
        35                  40                  45
Asn Ala Trp His Arg His Trp Ser Arg Trp Ile Leu Cys Glu Leu Gly
    50                  55                  60
Leu Leu Asn Gln Pro Val Leu Ser Leu Asp Pro Pro Gln Leu Lys Val
65                  70                  75                  80
Ala Leu Leu Ser Thr Asp Ala Leu Arg Thr Cys Ala Ala His Ala Gly
                85                  90                  95
Ala Leu Leu Cys Ala Pro Arg Leu Arg Arg Ala Ile Asp Gly Ala Glu
            100                 105                 110
Val Arg Thr Leu His Ala Ala Leu Gly Arg Asp Val Met Asn Phe Ala
        115                 120                 125
Val Ser Ser Ala Ala Arg Ala Leu His Asp Gly Leu Ala Ala Ser Ser
    130                 135                 140
Asp Trp Thr Leu Ala Ala Thr Val Gln Ala Ala Gln Lys Leu Gly Trp
145                 150                 155                 160
```

```
Ala Leu Leu Arg Asp Ala Val Gln Gly Ala Ala Asp Glu Ile Ala Leu
                165                 170                 175
Arg Cys Ala Leu Lys Leu Pro Arg Asp Leu Asp Pro Ala Pro Val Leu
            180                 185                 190
Pro Pro Glu Ala Ala Leu Ala Leu Val Leu Ser Met Leu Glu Ile Leu
        195                 200                 205
Asp Ala Glu Trp Leu Ser Ser Phe Pro Ala Gln Ala
    210                 215                 220
```

<210> 19
<211> 639
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(639)

<400> 19

```
atg gct ttc ctc gtt ccc cgc cca agc ctg atc cag gcg gta cgg ccc        48
Met ala Phe Leu Val Pro Arg Pro Ser Leu Ile Gln Ala Val Arg Pro
 1               5                  10                  15

ggc cgt gcg gat ccc gcg acc gac gtc ttg cgc gct gaa gac tac gcc        96
Gly Arg Ala Asp Pro Ala Thr Asp Val Leu Arg Ala Glu Asp Tyr Ala
            20                  25                  30

gag ctg ctc agc gcc gcg cag atc gtt gcc cag gca cac cgg cgg gcc       144
Glu Leu Leu Ser Ala Ala Gln Ile Val Ala Gln Ala His Arg Arg Ala
        35                  40                  45

ggc gaa atc gtg gcc gag gcg cga gag gag ttc gag cgc gag cgc agg       192
Gly Glu Ile Val Ala Glu Ala Arg Glu Glu Phe Glu Arg Glu Arg Arg
    50                  55                  60

cga ggc tat gag gag ggg cgc cgc gaa gcg ctt acg gat cag gcg gag       240
Arg Gly Tyr Glu Glu Gly Arg Arg Glu Ala Leu Thr Asp Gln Ala Glu
65                  70                  75                  80

aag atg ata gaa acc gta agc cgc acg atc gac tac ttc gcg ggt atc       288
Lys Met Ile Glu Thr Val Ser Arg Thr Ile Asp Tyr Phe Ala Gly Ile
                85                  90                  95

gag aac gag atg atc gaa ctg gtc atg agt gcg gtc cgc aag atc gtc       336
Glu Asn Glu Met Ile Glu Leu Val Met Ser Ala Val Arg Lys Ile Val
            100                 105                 110

gac ggt tac gac gac cgc gag cgc acc gtg atc gcc gtg cgc aac gca       384
Asp Gly Tyr Asp Asp Arg Glu Arg Thr Val Ile Ala Val Arg Asn Ala
        115                 120                 125

ttg gcg gtc gtg cgc aat cag cgc cag atg acc ttg cgc ctg cac cca       432
Leu Ala Val Val Arg Asn Gln Arg Gln Met Thr Leu Arg Leu His Pro
    130                 135                 140

gac gag gtg gat gtg ctc cgg gaa ggc atg aac cag ctt ctg gcg gcc       480
Asp Glu Val Asp Val Leu Arg Glu Gly Met Asn Gln Leu Leu Ala Ala
145                 150                 155                 160

tat ccg ggc gtg ggc tac ctg gac ctg ctg ccc gac gcc agg ctg gcg       528
```

```
Tyr Pro Gly Val Gly Tyr Leu Asp Leu Leu Pro Asp Ala Arg Leu Ala
                165                 170                 175

ccg gga gcc tgc ata ctg gag agc gag ata ggc atg gtc gag gcc agc      576
Pro Gly Ala Cys Ile Leu Glu Ser Glu Ile Gly Met Val Glu Ala Ser
            180                 185                 190

ctc gag gac cag ctg tgc gcc ttg cgg gcg gcc ttc gaa cgt aca ttc      624
Leu Glu Asp Gln Leu Cys Ala Leu Arg Ala Ala Phe Glu Arg Thr Phe
        195                 200                 205

ggc cgg cgc gga tag                                                  639
Gly Arg Arg Gly  *
    210
```

<210> 20
<211> 212
<212> PRT
<213> Bordetella pertussis

<400> 20

```
Met ala Phe Leu Val Pro Arg Pro Ser Leu Ile Gln Ala Val Arg Pro
 1               5                  10                  15
Gly Arg Ala Asp Pro Ala Thr Asp Val Leu Arg Ala Glu Asp Tyr Ala
            20                  25                  30
Glu Leu Leu Ser Ala Ala Gln Ile Val Ala Gln Ala His Arg Arg Ala
        35                  40                  45
Gly Glu Ile Val Ala Glu Ala Arg Glu Glu Phe Glu Arg Glu Arg Arg
    50                  55                  60
Arg Gly Tyr Glu Glu Gly Arg Arg Glu Ala Leu Thr Asp Gln Ala Glu
65                  70                  75                  80
Lys Met Ile Glu Thr Val Ser Arg Thr Ile Asp Tyr Phe Ala Gly Ile
                85                  90                  95
Glu Asn Glu Met Ile Glu Leu Val Met Ser Ala Val Arg Lys Ile Val
            100                 105                 110
Asp Gly Tyr Asp Asp Arg Glu Arg Thr Val Ile Ala Val Arg Asn Ala
        115                 120                 125
Leu Ala Val Val Arg Asn Gln Arg Gln Met Thr Leu Arg Leu His Pro
    130                 135                 140
Asp Glu Val Asp Val Leu Arg Glu Gly Met Asn Gln Leu Leu Ala Ala
145                 150                 155                 160
Tyr Pro Gly Val Gly Tyr Leu Asp Leu Leu Pro Asp Ala Arg Leu Ala
                165                 170                 175
Pro Gly Ala Cys Ile Leu Glu Ser Glu Ile Gly Met Val Glu Ala Ser
            180                 185                 190
Leu Glu Asp Gln Leu Cys Ala Leu Arg Ala Ala Phe Glu Arg Thr Phe
        195                 200                 205
Gly Arg Arg Gly
    210
```

<210> 21
<211> 1335
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1335)

<400> 21

```
atg cgt cag tac cac tac atc acg gag atg atg cgg gtg gcc ctg cag      48
Met Arg Gln Tyr His Tyr Ile Thr Glu Met Met Arg Val Ala Leu Gln
 1               5                  10                  15

gat ctg tcc acg ctg cgg ata aag ggc cgg gtg gtg caa gtg gtg gga      96
Asp Leu Ser Thr Leu Arg Ile Lys Gly Arg Val Val Gln Val Val Gly
            20                  25                  30

acg atc atc aag gcc gtc gtt ccg atg gtc aag atc ggc gaa gtg tgc     144
Thr Ile Ile Lys Ala Val Val Pro Met Val Lys Ile Gly Glu Val Cys
        35                  40                  45

ctg ctg cgc aat ccc ggc gag gac ttc gag atg cac ggc gaa gtg gtg     192
Leu Leu Arg Asn Pro Gly Glu Asp Phe Glu Met His Gly Glu Val Val
    50                  55                  60

ggc ttt gtc cgc gac gcc gcc ttg ctc acg cct atc ggc gac atg tac     240
Gly Phe Val Arg Asp Ala Ala Leu Leu Thr Pro Ile Gly Asp Met Tyr
65                  70                  75                  80

ggg att tcc tcg gcg acc gag gtg ata ccg acc gga cgc acg cat atg     288
Gly Ile Ser Ser Ala Thr Glu Val Ile Pro Thr Gly Arg Thr His Met
                85                  90                  95

gtc ccc gtc ggt ccg ggc ttg ctg gga cgc gtg ctg gac ggg ctg gga     336
Val Pro Val Gly Pro Gly Leu Leu Gly Arg Val Leu Asp Gly Leu Gly
            100                 105                 110

cgt ccg ctg gac gcc gcc gag tca ggg ccg ctg cat gcc cac aag ttc     384
Arg Pro Leu Asp Ala Ala Glu Ser Gly Pro Leu His Ala His Lys Phe
        115                 120                 125

tat ccg gtc ttc gcc gat gcg cca gac ccg ctg acg cgt cgc atc atc     432
Tyr Pro Val Phe Ala Asp Ala Pro Asp Pro Leu Thr Arg Arg Ile Ile
    130                 135                 140

cat gct ccg ctg gag ctg ggg gtg cgc gta ctg gac ggt ttg ctt aca     480
His Ala Pro Leu Glu Leu Gly Val Arg Val Leu Asp Gly Leu Leu Thr
145                 150                 155                 160

tgc ggg gaa ggc cag cgt ctg gga att ttc gca gcc gcc ggc ggc ggc     528
Cys Gly Glu Gly Gln Arg Leu Gly Ile Phe Ala Ala Ala Gly Gly Gly
                165                 170                 175

aag tcg acc ctg ctg ggc atg ctg gtc aag ggc gcc gcg gtc gac gtg     576
Lys Ser Thr Leu Leu Gly Met Leu Val Lys Gly Ala Ala Val Asp Val
            180                 185                 190

acg gtg gtg gcg ctg atc ggc gag cgt ggg cgg gaa gtt cgc gag ttc     624
Thr Val Val Ala Leu Ile Gly Glu Arg Gly Arg Glu Val Arg Glu Phe
        195                 200                 205

ctt gag cac gaa ctc ggt ccg gag ggc aga cgc aag agc gtg atc gtc     672
Leu Glu His Glu Leu Gly Pro Glu Gly Arg Arg Lys Ser Val Ile Val
    210                 215                 220

tgc gcg acc agc gac aag tcc tcg atg gag cgt gcc aag gcg gcg tac     720
Cys Ala Thr Ser Asp Lys Ser Ser Met Glu Arg Ala Lys Ala Ala Tyr
225                 230                 235                 240
```

```
gtc gca acc gcc atc gcc gaa tac ttc cgc gat caa ggg cag cgt gta      768
Val Ala Thr Ala Ile Ala Glu Tyr Phe Arg Asp Gln Gly Gln Arg Val
                245                 250                 255

ctt ttt ctg atg gac tcg gtc acc cgc ttt gcg cga gcc cag cgt gaa      816
Leu Phe Leu Met Asp Ser Val Thr Arg Phe Ala Arg Ala Gln Arg Glu
            260                 265                 270

atc ggc ttg gcg gca ggc gag ccg ccg acg cgg cgc ggc tat cca ccg      864
Ile Gly Leu Ala Ala Gly Glu Pro Pro Thr Arg Arg Gly Tyr Pro Pro
        275                 280                 285

tcg gtg ttc gcc acc ttg ccc aaa ctg atg gag cgc gcc ggc atg aac      912
Ser Val Phe Ala Thr Leu Pro Lys Leu Met Glu Arg Ala Gly Met Asn
    290                 295                 300

cag acg ggt tcg atc acg gcg ctg tat acg gtg ctg gtc gag ggg gac      960
Gln Thr Gly Ser Ile Thr Ala Leu Tyr Thr Val Leu Val Glu Gly Asp
305                 310                 315                 320

gac atg aac gaa ccg gtg gcc gac gag acg cgt tcg ata ctg gac ggc     1008
Asp Met Asn Glu Pro Val Ala Asp Glu Thr Arg Ser Ile Leu Asp Gly
                325                 330                 335

cac atc gtg ctc tcg cgc aag ctg gga gcg gcg aat cac tat cct gcc     1056
His Ile Val Leu Ser Arg Lys Leu Gly Ala Ala Asn His Tyr Pro Ala
            340                 345                 350

gtc gac gtg ctg gcc tcg gcc agc cgg gtc atg aat gcc gtg gtg tcg     1104
Val Asp Val Leu Ala Ser Ala Ser Arg Val Met Asn Ala Val Val Ser
        355                 360                 365

ccg cgt cac aag tac ctg gcc gga cgt atg cgc gaa ctg atg gcc aag     1152
Pro Arg His Lys Tyr Leu Ala Gly Arg Met Arg Glu Leu Met ala Lys
    370                 375                 380

tac cag gat gtc gag ctg ttg gtg aaa atc ggc gag tac aag cag ggc     1200
Tyr Gln Asp Val Glu Leu Leu Val Lys Ile Gly Glu Tyr Lys Gln Gly
385                 390                 395                 400

gcc gat gcg tcg acc gat gag gcg ata cag aag atc gga cag atc aat     1248
Ala Asp Ala Ser Thr Asp Glu Ala Ile Gln Lys Ile Gly Gln Ile Asn
                405                 410                 415

gcg ttt ctc aga caa cta acc gac gaa cgc gaa gca ttc gag gat acc     1296
Ala Phe Leu Arg Gln Leu Thr Asp Glu Arg Glu Ala Phe Glu Asp Thr
            420                 425                 430

gta ctg cgc atg gct gaa atc atc gga ccc gaa tcc taa                 1335
Val Leu Arg Met ala Glu Ile Ile Gly Pro Glu Ser  *
        435                 440
```

<210> 22
<211> 444
<212> PRT
<213> Bordetella pertussis

<400> 22

Met Arg Gln Tyr His Tyr Ile Thr Glu Met Met Arg Val Ala Leu Gln
1 5 10 15
Asp Leu Ser Thr Leu Arg Ile Lys Gly Arg Val Val Gln Val Val Gly
20 25 30
Thr Ile Ile Lys Ala Val Val Pro Met Val Lys Ile Gly Glu Val Cys
35 40 45
Leu Leu Arg Asn Pro Gly Glu Asp Phe Glu Met His Gly Glu Val Val
50 55 60
Gly Phe Val Arg Asp Ala Ala Leu Leu Thr Pro Ile Gly Asp Met Tyr
65 70 75 80
Gly Ile Ser Ser Ala Thr Glu Val Ile Pro Thr Gly Arg Thr His Met
85 90 95
Val Pro Val Gly Pro Gly Leu Leu Gly Arg Val Leu Asp Gly Leu Gly
100 105 110
Arg Pro Leu Asp Ala Ala Glu Ser Gly Pro Leu His Ala His Lys Phe
115 120 125
Tyr Pro Val Phe Ala Asp Ala Pro Asp Pro Leu Thr Arg Arg Ile Ile
130 135 140
His Ala Pro Leu Glu Leu Gly Val Arg Val Leu Asp Gly Leu Leu Thr
145 150 155 160
Cys Gly Glu Gly Gln Arg Leu Gly Ile Phe Ala Ala Ala Gly Gly Gly
165 170 175
Lys Ser Thr Leu Leu Gly Met Leu Val Lys Gly Ala Ala Val Asp Val
180 185 190
Thr Val Val Ala Leu Ile Gly Glu Arg Gly Arg Glu Val Arg Glu Phe
195 200 205
Leu Glu His Glu Leu Gly Pro Glu Gly Arg Arg Lys Ser Val Ile Val
210 215 220
Cys Ala Thr Ser Asp Lys Ser Ser Met Glu Arg Ala Lys Ala Ala Tyr
225 230 235 240
Val Ala Thr Ala Ile Ala Glu Tyr Phe Arg Asp Gln Gly Gln Arg Val
245 250 255
Leu Phe Leu Met Asp Ser Val Thr Arg Phe Ala Arg Ala Gln Arg Glu
260 265 270
Ile Gly Leu Ala Ala Gly Glu Pro Pro Thr Arg Arg Gly Tyr Pro Pro
275 280 285
Ser Val Phe Ala Thr Leu Pro Lys Leu Met Glu Arg Ala Gly Met Asn
290 295 300
Gln Thr Gly Ser Ile Thr Ala Leu Tyr Thr Val Leu Val Glu Gly Asp
305 310 315 320
Asp Met Asn Glu Pro Val Ala Asp Glu Thr Arg Ser Ile Leu Asp Gly
325 330 335
His Ile Val Leu Ser Arg Lys Leu Gly Ala Ala Asn His Tyr Pro Ala
340 345 350
Val Asp Val Leu Ala Ser Ala Ser Arg Val Met Asn Ala Val Val Ser
355 360 365
Pro Arg His Lys Tyr Leu Ala Gly Arg Met Arg Glu Leu Met *ala* Lys
370 375 380
Tyr Gln Asp Val Glu Leu Leu Val Lys Ile Gly Glu Tyr Lys Gln Gly
385 390 395 400
Ala Asp Ala Ser Thr Asp Glu Ala Ile Gln Lys Ile Gly Gln Ile Asn
405 410 415
Ala Phe Leu Arg Gln Leu Thr Asp Glu Arg Glu Ala Phe Glu Asp Thr
420 425 430
Val Leu Arg Met *ala* Glu Ile Ile Gly Pro Glu Ser
435 440

<210> 23
<211> 510
<212> DNA

<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(510)

<400> 23

```
atg gac ctg gaa agc ctg ctt gcc atc aag cat ttt cgc gcc gac caa        48
Met Asp Leu Glu Ser Leu Leu Ala Ile Lys His Phe Arg Ala Asp Gln
 1               5                  10                  15

gcc cag ctt gcg ctg aaa cgc caa cag cag gcc tgc gcg gtt gct gcc        96
Ala Gln Leu Ala Leu Lys Arg Gln Gln Ala Cys Ala Val Ala Ala
            20                  25                  30

gcg gcg cag cgt cag gcg caa ggc cgc ctc gac gat tgt cgc ctg tgg       144
Ala Ala Gln Arg Gln Ala Gln Gly Arg Leu Asp Asp Cys Arg Leu Trp
        35                  40                  45

gcc gga cag ctc gaa aac cgt cta tat gcc gag ctg tgc cgg cgc atc       192
Ala Gly Gln Leu Glu Asn Arg Leu Tyr Ala Glu Leu Cys Arg Arg Ile
    50                  55                  60

gtc aag aca cgc gac atc gac gag gtg ctg caa cga gtg ggc cac gcc       240
Val Lys Thr Arg Asp Ile Asp Glu Val Leu Gln Arg Val Gly His Ala
65                  70                  75                  80

cgc gac cgc cag gcc agc ctg gcg ctg cag ctc gac gac gcc gtg cgc       288
Arg Asp Arg Gln Ala Ser Leu Ala Leu Gln Leu Asp Asp Ala Val Arg
                85                  90                  95

cgt cac gaa cat gaa atc cag ctg ctc gcg cag cag cgc gag cag cac       336
Arg His Glu His Glu Ile Gln Leu Leu Ala Gln Gln Arg Glu Gln His
            100                 105                 110

cgg gag tgc ttc cag gcg cag caa cgg atc gcc gag ttg gtg cgc ctg       384
Arg Glu Cys Phe Gln Ala Gln Gln Arg Ile Ala Glu Leu Val Arg Leu
        115                 120                 125

cag cag gtc gag gcg gcg gcc ttg cgc gag agc cag gaa gat cgc gaa       432
Gln Gln Val Glu Ala Ala Ala Leu Arg Glu Ser Gln Glu Asp Arg Glu
    130                 135                 140

att cag gaa gcc atc gaa ttg tcg gcg cgt ggg cgc gac gat gca tcg       480
Ile Gln Glu Ala Ile Glu Leu Ser Ala Arg Gly Arg Asp Asp Ala Ser
145                 150                 155                 160

cga gcc ggc gac ggc ctg gcg cgg cta tga                               510
Arg Ala Gly Asp Gly Leu Ala Arg Leu  *
                165
```

<210> 24
<211> 169
<212> PRT
<213> Bordetella pertussis

<400> 24

```
Met Asp Leu Glu Ser Leu Leu Ala Ile Lys His Phe Arg Ala Asp Gln
 1               5                  10                  15
```

```
Ala Gln Leu Ala Leu Lys Arg Gln Gln Gln Ala Cys Ala Val Ala Ala
            20                  25                  30
Ala Ala Gln Arg Gln Ala Gln Gly Arg Leu Asp Asp Cys Arg Leu Trp
        35                  40                  45
Ala Gly Gln Leu Glu Asn Arg Leu Tyr Ala Glu Leu Cys Arg Arg Ile
    50                  55                  60
Val Lys Thr Arg Asp Ile Asp Glu Val Leu Gln Arg Val Gly His Ala
65                  70                  75                  80
Arg Asp Arg Gln Ala Ser Leu Ala Leu Gln Leu Asp Asp Ala Val Arg
                85                  90                  95
Arg His Glu His Glu Ile Gln Leu Leu Ala Gln Gln Arg Glu Gln His
            100                 105                 110
Arg Glu Cys Phe Gln Ala Gln Gln Arg Ile Ala Glu Leu Val Arg Leu
        115                 120                 125
Gln Gln Val Glu Ala Ala Ala Leu Arg Glu Ser Gln Glu Asp Arg Glu
    130                 135                 140
Ile Gln Glu Ala Ile Glu Leu Ser Ala Arg Gly Arg Asp Asp Ala Ser
145                 150                 155                 160
Arg Ala Gly Asp Gly Leu Ala Arg Leu
                165
```

<210> 25
<211> 549
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(549)

<400> 25

```
atg aac cag cca gac ggg ctg ggt tcg ccc atg gcc ggc ggc ggg cag      48
Met Asn Gln Pro Asp Gly Leu Gly Ser Pro Met ala Gly Gly Gly Gln
 1               5                  10                  15

cgc atg ggc gtg gcg cgc acg ccg tat gcg cgt cag ccg gat cgg gat      96
Arg Met Gly Val Ala Arg Thr Pro Tyr Ala Arg Gln Pro Asp Arg Asp
            20                  25                  30

gcg cag cgt gcc ttc gag cgg gaa atg gaa cag gag aaa gcg aag gaa     144
Ala Gln Arg Ala Phe Glu Arg Glu Met Glu Gln Glu Lys Ala Lys Glu
        35                  40                  45

gaa ctg ccc ggg ccg caa cgc ctg gcg ccg ggt ccg gcc tgc gtc ggc     192
Glu Leu Pro Gly Pro Gln Arg Leu Ala Pro Gly Pro Ala Cys Val Gly
    50                  55                  60

tgg ctg gcg tcg atg gaa cct gcc gcc ggc cgt cca ccg gcc agt ctg     240
Trp Leu Ala Ser Met Glu Pro Ala Ala Gly Arg Pro Pro Ala Ser Leu
65                  70                  75                  80

gcc cag gcg ctg gca agc gtg gct gcg ggg ctg gcg gta ggc gac gtg     288
Ala Gln Ala Leu Ala Ser Val Ala Ala Gly Leu Ala Val Gly Asp Val
                85                  90                  95

ctg gag ggg tat cgc gaa gcc cgt atc gtt gtg gac gat acg ctg cta     336
Leu Glu Gly Tyr Arg Glu Ala Arg Ile Val Val Asp Asp Thr Leu Leu
            100                 105                 110

ccc gac acc acc ttg tcg gta cgg gag gac ggc ggc tgg atc gtg gtg     384
```

```
Pro Asp Thr Thr Leu Ser Val Arg Glu Asp Gly Gly Trp Ile Val Val
        115                 120                 125

gct ttc gca tgc cga caa cgg gac gct tgc gag cgc ctg cac gcg tgc      432
Ala Phe Ala Cys Arg Gln Arg Asp Ala Cys Glu Arg Leu His Ala Cys
    130                 135                 140

gcc gac cgg ttg gcc atg gag ctc gcg ctg gag ctg gcg cgc gac gtc      480
Ala Asp Arg Leu Ala Met Glu Leu Ala Leu Glu Leu Ala Arg Asp Val
145                 150                 155                 160

gag gtt gcg gtg gca tgc gac ggc gag ccg cac gag cgg gtg gcg cgc      528
Glu Val Ala Val Ala Cys Asp Gly Glu Pro His Glu Arg Val Ala Arg
                165                 170                 175

gcg cag cgg ccg tgg cga tga                                         549
Ala Gln Arg Pro Trp Arg  *
            180
```

<210> 26
<211> 182
<212> PRT
<213> Bordetella pertussis

<400> 26

```
Met Asn Gln Pro Asp Gly Leu Gly Ser Pro Met ala Gly Gly Gly Gln
 1               5                  10                  15
Arg Met Gly Val Ala Arg Thr Pro Tyr Ala Arg Gln Pro Asp Arg Asp
            20                  25                  30
Ala Gln Arg Ala Phe Glu Arg Glu Met Glu Gln Glu Lys Ala Lys Glu
        35                  40                  45
Glu Leu Pro Gly Pro Gln Arg Leu Ala Pro Gly Pro Ala Cys Val Gly
    50                  55                  60
Trp Leu Ala Ser Met Glu Pro Ala Ala Gly Arg Pro Pro Ala Ser Leu
65                  70                  75                  80
Ala Gln Ala Leu Ala Ser Val Ala Ala Gly Leu Ala Val Gly Asp Val
                85                  90                  95
Leu Glu Gly Tyr Arg Glu Ala Arg Ile Val Val Asp Asp Thr Leu Leu
            100                 105                 110
Pro Asp Thr Thr Leu Ser Val Arg Glu Asp Gly Gly Trp Ile Val Val
        115                 120                 125
Ala Phe Ala Cys Arg Gln Arg Asp Ala Cys Glu Arg Leu His Ala Cys
    130                 135                 140
Ala Asp Arg Leu Ala Met Glu Leu Ala Leu Glu Leu Ala Arg Asp Val
145                 150                 155                 160
Glu Val Ala Val Ala Cys Asp Gly Glu Pro His Glu Arg Val Ala Arg
                165                 170                 175
Ala Gln Arg Pro Trp Arg
            180
```

<210> 27
<211> 1080
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1080)

<400> 27

```
atg aat cga gtg gcc ggc ggg gcg gcg gcg cag gcc gct ggc atg gtg        48
Met Asn Arg Val Ala Gly Gly Ala Ala Ala Gln Ala Ala Gly Met Val
 1               5                  10                  15

gat ctc gcg gtt ccg cgg ttg agc gcc ggc gag gcc cat gcc ctg tcg        96
Asp Leu Ala Val Pro Arg Leu Ser Ala Gly Glu Ala His Ala Leu Ser
            20                  25                  30

agg att gca tgc cat ggc gcg cga ttc gac gtt cgg ctt ggc gag ccg       144
Arg Ile Ala Cys His Gly Ala Arg Phe Asp Val Arg Leu Gly Glu Pro
        35                  40                  45

gcc gtg cgc tgg cac tgc gcc ctg acg cct tgc gtg cac ggc gac ctt       192
Ala Val Arg Trp His Cys Ala Leu Thr Pro Cys Val His Gly Asp Leu
    50                  55                  60

gcc gat ggc gag atg gaa agc ctg caa ctg caa tgg gcc ggg acg tac       240
Ala Asp Gly Glu Met Glu Ser Leu Gln Leu Gln Trp Ala Gly Thr Tyr
65                  70                  75                  80

atc ggc ctg acg gtt ccg cgc gcg gcc gcg gcg gga tgg ctg gcg gcg       288
Ile Gly Leu Thr Val Pro Arg Ala Ala Ala Ala Gly Trp Leu Ala Ala
                85                  90                  95

cgc ctg ccc cgg ttt tcc ggc gtg gag ttg ccg gaa ccc att gcg gcg       336
Arg Leu Pro Arg Phe Ser Gly Val Glu Leu Pro Glu Pro Ile Ala Ala
            100                 105                 110

gcg gcc ctg gag gca atg ctg gag gag gtc tgt cga ggc gtg gcc gga       384
Ala Ala Leu Glu Ala Met Leu Glu Glu Val Cys Arg Gly Val Ala Gly
        115                 120                 125

ctc gac cag caa ggc ccg gtc cgc gtg gcg cgg caa ggc ggg acg cca       432
Leu Asp Gln Gln Gly Pro Val Arg Val Ala Arg Gln Gly Gly Thr Pro
    130                 135                 140

ccg gtc cag ccg cat cgc tgg acc ctg acg gta cgg gcg cct gac ggt       480
Pro Val Gln Pro His Arg Trp Thr Leu Thr Val Arg Ala Pro Asp Gly
145                 150                 155                 160

ggc gtc tgg cgc gcg gta ctg gcg tgc gac gca tgg gcc ttg caa gcg       528
Gly Val Trp Arg Ala Val Leu Ala Cys Asp Ala Trp Ala Leu Gln Ala
                165                 170                 175

gtc gcg gcg gcg ctg gat tcc gtt gcg cct gcc gat ggt cgg gtc aat       576
Val Ala Ala Ala Leu Asp Ser Val Ala Pro Ala Asp Gly Arg Val Asn
            180                 185                 190

ccg gag cgc gtg ccg gtc agg ttg cgt gcc gat gtc ggc gcg gcg tcc       624
Pro Glu Arg Val Pro Val Arg Leu Arg Ala Asp Val Gly Ala Ala Ser
        195                 200                 205

gtg acc gca ggc cag ctg cgg acg ctg cga gcg ggc gac gtc gtg ttg       672
Val Thr Ala Gly Gln Leu Arg Thr Leu Arg Ala Gly Asp Val Val Leu
    210                 215                 220

ctc gcg cag tac cgg gtg agc gat gcc gca gaa cta tgg ttg tcg gcc       720
Leu Ala Gln Tyr Arg Val Ser Asp Ala Ala Glu Leu Trp Leu Ser Ala
225                 230                 235                 240
```

```
gga ccc agc gcg atc cgg gta cgg gcc gag cat gcg tct ttt cgt gta      768
Gly Pro Ser Ala Ile Arg Val Arg Ala Glu His Ala Ser Phe Arg Val
                245                 250                 255

act caa ggt tgg act ccc atc atg acg gaa ccc gcg aca cct gac cct      816
Thr Gln Gly Trp Thr Pro Ile Met Thr Glu Pro Ala Thr Pro Asp Pro
            260                 265                 270

ggc gaa acc ccg gca cag gcc gac gcg acg ctc gat acc gat cag ata      864
Gly Glu Thr Pro Ala Gln Ala Asp Ala Thr Leu Asp Thr Asp Gln Ile
        275                 280                 285

ccc gtg cgc ctg acg ttc gac ctg ggc gag cgc gag ttc acg ctt gcg      912
Pro Val Arg Leu Thr Phe Asp Leu Gly Glu Arg Glu Phe Thr Leu Ala
    290                 295                 300

cag ctg cgc agc ctg cat ccg ggc tgc acg ttc gac ctc gag cgg ccc      960
Gln Leu Arg Ser Leu His Pro Gly Cys Thr Phe Asp Leu Glu Arg Pro
305                 310                 315                 320

atc gcc gac ggg ccg gtc atg gtg cgg gcc aat ggc ctg ttg ctg ggc     1008
Ile Ala Asp Gly Pro Val Met Val Arg Ala Asn Gly Leu Leu Leu Gly
                325                 330                 335

agc ggc cgg ctg gtc gac atc gac ggc cgc atc ggc gtg gta ttg cag     1056
Ser Gly Arg Leu Val Asp Ile Asp Gly Arg Ile Gly Val Val Leu Gln
            340                 345                 350

tcg gtc agg cct gga ctc gca tga                                    1080
Ser Val Arg Pro Gly Leu Ala  *
        355
```

<210> 28
<211> 359
<212> PRT
<213> Bordetella pertussis

<400> 28

```
Met Asn Arg Val Ala Gly Gly Ala Ala Ala Gln Ala Ala Gly Met Val
 1               5                  10                  15
Asp Leu Ala Val Pro Arg Leu Ser Ala Gly Glu Ala His Ala Leu Ser
            20                  25                  30
Arg Ile Ala Cys His Gly Ala Arg Phe Asp Val Arg Leu Gly Glu Pro
        35                  40                  45
Ala Val Arg Trp His Cys Ala Leu Thr Pro Cys Val His Gly Asp Leu
    50                  55                  60
Ala Asp Gly Glu Met Glu Ser Leu Gln Leu Gln Trp Ala Gly Thr Tyr
65                  70                  75                  80
Ile Gly Leu Thr Val Pro Arg Ala Ala Ala Ala Gly Trp Leu Ala Ala
                85                  90                  95
Arg Leu Pro Arg Phe Ser Gly Val Glu Leu Pro Glu Pro Ile Ala Ala
            100                 105                 110
Ala Ala Leu Glu Ala Met Leu Glu Glu Val Cys Arg Gly Val Ala Gly
        115                 120                 125
Leu Asp Gln Gln Gly Pro Val Arg Val Ala Arg Gln Gly Gly Thr Pro
    130                 135                 140
Pro Val Gln Pro His Arg Trp Thr Leu Thr Val Arg Ala Pro Asp Gly
145                 150                 155                 160
```

```
Gly Val Trp Arg Ala Val Leu Ala Cys Asp Ala Trp Ala Leu Gln Ala
                165                 170                 175
Val Ala Ala Ala Leu Asp Ser Val Ala Pro Ala Asp Gly Arg Val Asn
            180                 185                 190
Pro Glu Arg Val Pro Val Arg Leu Arg Ala Asp Val Gly Ala Ala Ser
        195                 200                 205
Val Thr Ala Gly Gln Leu Arg Thr Leu Arg Ala Gly Asp Val Val Leu
    210                 215                 220
Leu Ala Gln Tyr Arg Val Ser Asp Ala Ala Glu Leu Trp Leu Ser Ala
225                 230                 235                 240
Gly Pro Ser Ala Ile Arg Val Arg Ala Glu His Ala Ser Phe Arg Val
                245                 250                 255
Thr Gln Gly Trp Thr Pro Ile Met Thr Glu Pro Ala Thr Pro Asp Pro
            260                 265                 270
Gly Glu Thr Pro Ala Gln Ala Asp Ala Thr Leu Asp Thr Asp Gln Ile
        275                 280                 285
Pro Val Arg Leu Thr Phe Asp Leu Gly Glu Arg Glu Phe Thr Leu Ala
    290                 295                 300
Gln Leu Arg Ser Leu His Pro Gly Cys Thr Phe Asp Leu Glu Arg Pro
305                 310                 315                 320
Ile Ala Asp Gly Pro Val Met Val Arg Ala Asn Gly Leu Leu Leu Gly
                325                 330                 335
Ser Gly Arg Leu Val Asp Ile Asp Gly Arg Ile Gly Val Val Leu Gln
            340                 345                 350
Ser Val Arg Pro Gly Leu Ala
        355
```

<210> 29
<211> 672
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(672)

<400> 29

```
atg agc gat acc gac ccc ttc agc ctg gcc ctg ttt ctg gcg ctg ctg      48
Met Ser Asp Thr Asp Pro Phe Ser Leu Ala Leu Phe Leu Ala Leu Leu
 1               5                  10                  15

gcg ctg gta ccg ctc atc gtc gtc atg acc acg tcg ttc ctg aag atc      96
Ala Leu Val Pro Leu Ile Val Val Met Thr Thr Ser Phe Leu Lys Ile
            20                  25                  30

gcc gtc gtg ctt gcc ttg gtg cgc aac gcc ctg gga gtg caa cag gta     144
Ala Val Val Leu Ala Leu Val Arg Asn Ala Leu Gly Val Gln Gln Val
        35                  40                  45

ccg ccc aac atg gcc ctg tac ggg ctg gcg ctt att ctt tcc gcg tat     192
Pro Pro Asn Met ala Leu Tyr Gly Leu Ala Leu Ile Leu Ser Ala Tyr
    50                  55                  60

gtg atg gcg ccg gtc gtt cac agg ata ggc acc gag gtc cag gcc ttg     240
Val Met ala Pro Val Val His Arg Ile Gly Thr Glu Val Gln Ala Leu
65                  70                  75                  80

acc gcg caa gcc ggg gag tcc ggc acc gcc gcg ccg atg gcg ctg gac     288
Thr Ala Gln Ala Gly Glu Ser Gly Thr Ala Ala Pro Met ala Leu Asp
                85                  90                  95
```

```
gcc gtg ctt ggc gtg gcc gag cga ggc gtg ggg ccg ctg cgg gcc ttc      336
Ala Val Leu Gly Val Ala Glu Arg Gly Val Gly Pro Leu Arg Ala Phe
            100                 105                 110

atg ttg cgc aac agc cag ccg gcc cag cgt gat ttc ttc ctg cgc aca      384
Met Leu Arg Asn Ser Gln Pro Ala Gln Arg Asp Phe Phe Leu Arg Thr
        115                 120                 125

gcg cgt cat ctc tgg ggc gag gag gca tcg cgg gac ctg tcg gaa gac      432
Ala Arg His Leu Trp Gly Glu Glu Ala Ser Arg Asp Leu Ser Glu Asp
    130                 135                 140

aac ctg ctg gta ttg acg ccc gca ttt ctg gtt tcg gag ctg acc gcc      480
Asn Leu Leu Val Leu Thr Pro Ala Phe Leu Val Ser Glu Leu Thr Ala
145                 150                 155                 160

gca ttc cag ctt ggc ttt ctg ctg tac ctg ccg ttc atc atc atc gac      528
Ala Phe Gln Leu Gly Phe Leu Leu Tyr Leu Pro Phe Ile Ile Ile Asp
                165                 170                 175

ctc atc gta tcg aac att ctt ctt gcc atg gga atg atg atg gtt tct      576
Leu Ile Val Ser Asn Ile Leu Leu Ala Met Gly Met Met Met Val Ser
            180                 185                 190

ccc gtg acg atc tcc atg ccg ttg aag ctg ttc ctg ttc gtc atg gtg      624
Pro Val Thr Ile Ser Met Pro Leu Lys Leu Phe Leu Phe Val Met Val
        195                 200                 205

gac ggc tgg acg cgc ctg atc cag ggc ctg gtg ctt tcc tat cgg tga      672
Asp Gly Trp Thr Arg Leu Ile Gln Gly Leu Val Leu Ser Tyr Arg  *
    210                 215                 220
```

<210> 30
<211> 223
<212> PRT
<213> Bordetella pertussis

<400> 30

```
Met Ser Asp Thr Asp Pro Phe Ser Leu Ala Leu Phe Leu Ala Leu Leu
 1               5                  10                  15
Ala Leu Val Pro Leu Ile Val Val Met Thr Thr Ser Phe Leu Lys Ile
            20                  25                  30
Ala Val Val Leu Ala Leu Val Arg Asn Ala Leu Gly Val Gln Gln Val
        35                  40                  45
Pro Pro Asn Met ala Leu Tyr Gly Leu Ala Leu Ile Leu Ser Ala Tyr
    50                  55                  60
Val Met ala Pro Val Val His Arg Ile Gly Thr Glu Val Gln Ala Leu
65                  70                  75                  80
Thr Ala Gln Ala Gly Glu Ser Gly Thr Ala Ala Pro Met ala Leu Asp
                85                  90                  95
Ala Val Leu Gly Val Ala Glu Arg Gly Val Gly Pro Leu Arg Ala Phe
            100                 105                 110
Met Leu Arg Asn Ser Gln Pro Ala Gln Arg Asp Phe Phe Leu Arg Thr
        115                 120                 125
Ala Arg His Leu Trp Gly Glu Glu Ala Ser Arg Asp Leu Ser Glu Asp
    130                 135                 140
Asn Leu Leu Val Leu Thr Pro Ala Phe Leu Val Ser Glu Leu Thr Ala
145                 150                 155                 160
```

```
Ala Phe Gln Leu Gly Phe Leu Leu Tyr Leu Pro Phe Ile Ile Ile Asp
                165                 170                 175
Leu Ile Val Ser Asn Ile Leu Leu Ala Met Gly Met Met Met Val Ser
            180                 185                 190
Pro Val Thr Ile Ser Met Pro Leu Lys Leu Phe Leu Phe Val Met Val
        195                 200                 205
Asp Gly Trp Thr Arg Leu Ile Gln Gly Leu Val Leu Ser Tyr Arg
    210                 215                 220
```

<210> 31
<211> 267
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(267)

<400> 31

```
atg caa acc caa gac ctg gtt tcg ttc atg aca cag gcg ttg tac ctg        48
Met Gln Thr Gln Asp Leu Val Ser Phe Met Thr Gln Ala Leu Tyr Leu
 1               5                  10                  15

gtg ctc tgg ctg tcg ctg ccg ccc atc gcc gtg gtg gcg atc gtg gga        96
Val Leu Trp Leu Ser Leu Pro Pro Ile Ala Val Val Ala Ile Val Gly
            20                  25                  30

acg ctg ttt tcc ctg ttg cag gcc ttg acg cag gtg cag gag cag acc       144
Thr Leu Phe Ser Leu Leu Gln Ala Leu Thr Gln Val Gln Glu Gln Thr
        35                  40                  45

ctg tcc ttc gcc gtg aag ctg ata gcc gtg ttc gcc acg ctg atg ctg       192
Leu Ser Phe Ala Val Lys Leu Ile Ala Val Phe Ala Thr Leu Met Leu
    50                  55                  60

gcg gcc cgg tgg ata agc gcg gaa atc tat aac ttc acg att gcg gtg       240
Ala Ala Arg Trp Ile Ser Ala Glu Ile Tyr Asn Phe Thr Ile Ala Val
65                  70                  75                  80

ttc gat gcc ttt cat cgg atc cac tga                                  267
Phe Asp Ala Phe His Arg Ile His  *
                85
```

<210> 32
<211> 88
<212> PRT
<213> Bordetella pertussis

<400> 32

```
Met Gln Thr Gln Asp Leu Val Ser Phe Met Thr Gln Ala Leu Tyr Leu
 1               5                  10                  15
Val Leu Trp Leu Ser Leu Pro Pro Ile Ala Val Val Ala Ile Val Gly
            20                  25                  30
Thr Leu Phe Ser Leu Leu Gln Ala Leu Thr Gln Val Gln Glu Gln Thr
        35                  40                  45
Leu Ser Phe Ala Val Lys Leu Ile Ala Val Phe Ala Thr Leu Met Leu
    50                  55                  60
Ala Ala Arg Trp Ile Ser Ala Glu Ile Tyr Asn Phe Thr Ile Ala Val
65                  70                  75                  80
```

```
Phe Asp Ala Phe His Arg Ile His
                85

      <210> 33
      <211> 801
      <212> DNA
      <213> Bordetella pertussis

      <220>
      <221> CDS
      <222> (1)...(801)

      <400> 33
atg cac acg gag ttc aat ttc gtc gag gcg aag gtt ttc ctg gga acg      48
Met His Thr Glu Phe Asn Phe Val Glu Ala Lys Val Phe Leu Gly Thr
 1               5                  10                  15

ctg gcc atg acg caa ccg cgg ata ctc acg gcc atg ctc ttt ctg ccg      96
Leu Ala Met Thr Gln Pro Arg Ile Leu Thr Ala Met Leu Phe Leu Pro
            20                  25                  30

atg ttc aac cgt cag ttt ctg cct ggt ccg ctg cgt tac gcc gtc ggc     144
Met Phe Asn Arg Gln Phe Leu Pro Gly Pro Leu Arg Tyr Ala Val Gly
        35                  40                  45

gcc tgt ctc ggg ctg atc gtg gtt ccc cag ctg gcg ccg cag tat gcc     192
Ala Cys Leu Gly Leu Ile Val Val Pro Gln Leu Ala Pro Gln Tyr Ala
    50                  55                  60

gcg ctg gat atc gac tgg ccc cgg ctg ctg gcg ctg ctg gcc aag gag     240
Ala Leu Asp Ile Asp Trp Pro Arg Leu Leu Ala Leu Leu Ala Lys Glu
65                  70                  75                  80

gcg atg gtg ggc atg ttc ctg ggt tgg ctg gct gcc ttg cca ttc tgg     288
Ala Met Val Gly Met Phe Leu Gly Trp Leu Ala Ala Leu Pro Phe Trp
                85                  90                  95

atc ttc gag gcc atc ggc ttc gtc ata gac aac caa cgg ggc gcc agc     336
Ile Phe Glu Ala Ile Gly Phe Val Ile Asp Asn Gln Arg Gly Ala Ser
            100                 105                 110

ctg ggc gct atc ctc aac ccc gcc acg ggc aac gat tcg tcg ccc atg     384
Leu Gly Ala Ile Leu Asn Pro Ala Thr Gly Asn Asp Ser Ser Pro Met
        115                 120                 125

ggc att ctc ttc aat ctg gga ttc atg gtg ttc ttc ctg acg gcg ggc     432
Gly Ile Leu Phe Asn Leu Gly Phe Met Val Phe Phe Leu Thr Ala Gly
    130                 135                 140

gga ttc ggg ttg ttc gcc acg atg ctg tat gac agc ttc ggg ttg tgg     480
Gly Phe Gly Leu Phe Ala Thr Met Leu Tyr Asp Ser Phe Gly Leu Trp
145                 150                 155                 160

aac atc tgg gcg tgg tgg ccg tcc atg ccc gca cag ggc gcc gtg cgg     528
Asn Ile Trp Ala Trp Trp Pro Ser Met Pro Ala Gln Gly Ala Val Arg
                165                 170                 175

atg ctg gac cag ttc agt ggc ttt gcc gcg cgt gtc ctg ctg ctg gcc     576
Met Leu Asp Gln Phe Ser Gly Phe Ala Ala Arg Val Leu Leu Leu Ala
            180                 185                 190
```

```
tcg ccg gcc atc gtg gcc atg ttc ctg gcc gag ctg ggc ctg gcc ctg      624
Ser Pro Ala Ile Val Ala Met Phe Leu Ala Glu Leu Gly Leu Ala Leu
        195                 200                 205

atc agc cgc ttc gcg cct caa ctg cag gtg ttc ttc ctg gct ctg ccg      672
Ile Ser Arg Phe Ala Pro Gln Leu Gln Val Phe Phe Leu Ala Leu Pro
    210                 215                 220

gta aag agc gcg ctg gtg ctg ttc gtg ctg gtg ctg tac atg gca acg      720
Val Lys Ser Ala Leu Val Leu Phe Val Leu Val Leu Tyr Met ala Thr
225                 230                 235                 240

ttg ttc cag tat gca ggc gaa atc ctg ggt tct gtg ggc cgg atc gtg      768
Leu Phe Gln Tyr Ala Gly Glu Ile Leu Gly Ser Val Gly Arg Ile Val
                245                 250                 255

ccg ttc ctg cat tca gcg tgg ccc ggc cca tga                          801
Pro Phe Leu His Ser Ala Trp Pro Gly Pro  *
            260                 265
```

<210> 34
<211> 266
<212> PRT
<213> Bordetella pertussis

<400> 34

```
Met His Thr Glu Phe Asn Phe Val Glu Ala Lys Val Phe Leu Gly Thr
 1               5                  10                  15
Leu Ala Met Thr Gln Pro Arg Ile Leu Thr Ala Met Leu Phe Leu Pro
            20                  25                  30
Met Phe Asn Arg Gln Phe Leu Pro Gly Pro Leu Arg Tyr Ala Val Gly
        35                  40                  45
Ala Cys Leu Gly Leu Ile Val Val Pro Gln Leu Ala Pro Gln Tyr Ala
    50                  55                  60
Ala Leu Asp Ile Asp Trp Pro Arg Leu Leu Ala Leu Leu Ala Lys Glu
65                  70                  75                  80
Ala Met Val Gly Met Phe Leu Gly Trp Leu Ala Ala Leu Pro Phe Trp
                85                  90                  95
Ile Phe Glu Ala Ile Gly Phe Val Ile Asp Asn Gln Arg Gly Ala Ser
            100                 105                 110
Leu Gly Ala Ile Leu Asn Pro Ala Thr Gly Asn Asp Ser Ser Pro Met
        115                 120                 125
Gly Ile Leu Phe Asn Leu Gly Phe Met Val Phe Phe Leu Thr Ala Gly
    130                 135                 140
Gly Phe Gly Leu Phe Ala Thr Met Leu Tyr Asp Ser Phe Gly Leu Trp
145                 150                 155                 160
Asn Ile Trp Ala Trp Trp Pro Ser Met Pro Ala Gln Gly Ala Val Arg
                165                 170                 175
Met Leu Asp Gln Phe Ser Gly Phe Ala Ala Arg Val Leu Leu Leu Ala
            180                 185                 190
Ser Pro Ala Ile Val Ala Met Phe Leu Ala Glu Leu Gly Leu Ala Leu
        195                 200                 205
Ile Ser Arg Phe Ala Pro Gln Leu Gln Val Phe Phe Leu Ala Leu Pro
    210                 215                 220
Val Lys Ser Ala Leu Val Leu Phe Val Leu Val Leu Tyr Met ala Thr
225                 230                 235                 240
Leu Phe Gln Tyr Ala Gly Glu Ile Leu Gly Ser Val Gly Arg Ile Val
                245                 250                 255
```

```
Pro Phe Leu His Ser Ala Trp Pro Gly Pro
            260                 265

      <210> 35
      <211> 1050
      <212> DNA
      <213> Bordetella pertussis

      <220>
      <221> CDS
      <222> (1)...(1050)

      <400> 35
atg agc ggc gag aaa acc gag cgg ccc acc ccg aag cgc ctg cgc gat      48
Met Ser Gly Glu Lys Thr Glu Arg Pro Thr Pro Lys Arg Leu Arg Asp
 1               5                  10                  15

tcc cgc gag aaa ggc gag gtc gca cac agc cgg gac ttt acc cag acg      96
Ser Arg Glu Lys Gly Glu Val Ala His Ser Arg Asp Phe Thr Gln Thr
            20                  25                  30

gcg ctg ata tgc gcc ttg ttc ggg cac ttt ctg atc aat gcc ccg tcc     144
Ala Leu Ile Cys Ala Leu Phe Gly His Phe Leu Ile Asn Ala Pro Ser
        35                  40                  45

att ctc gcg tcg ctg cga gcg ctg ata ctg gcg ccg gcg gcc ttt gcc     192
Ile Leu Ala Ser Leu Arg Ala Leu Ile Leu Ala Pro Ala Ala Phe Ala
    50                  55                  60

gac cag ggg ttc gcc gtc gca ttg ggg ccc gtg ctg acg gaa atc ctc     240
Asp Gln Gly Phe Ala Val Ala Leu Gly Pro Val Leu Thr Glu Ile Leu
65                  70                  75                  80

gat cag gcc gtc cgc gtg ctc gct ccg ctg att ctc atc gtg ctt ggg     288
Asp Gln Ala Val Arg Val Leu Ala Pro Leu Ile Leu Ile Val Leu Gly
                85                  90                  95

gtg ggg atg ttc gcc gaa ttc ctg cag gta ggc gtc gtg ctg gcg ttt     336
Val Gly Met Phe Ala Glu Phe Leu Gln Val Gly Val Val Leu Ala Phe
            100                 105                 110

cga aag ctc aag cct tcg gcg gag aaa ctg aat ccc gcc ggc aat ttg     384
Arg Lys Leu Lys Pro Ser Ala Glu Lys Leu Asn Pro Ala Gly Asn Leu
        115                 120                 125

aag aat atc ttc tcg gcg cgc aac ctg atg gag ttc atc aag tcg gta     432
Lys Asn Ile Phe Ser Ala Arg Asn Leu Met Glu Phe Ile Lys Ser Val
    130                 135                 140

tgc aag atc ctg ttt ctg gcg gtg ttg gtc acg ttg gtg ata cgg gat     480
Cys Lys Ile Leu Phe Leu Ala Val Leu Val Thr Leu Val Ile Arg Asp
145                 150                 155                 160

tcc ttg cag ccg ctg atg gcc gtt ccc cat agc ggg ctg gac ggg ttg     528
Ser Leu Gln Pro Leu Met ala Val Pro His Ser Gly Leu Asp Gly Leu
                165                 170                 175

cga acg ggc gta ggc cgc att ctg cag gtc atg gtc tgg aac atc gga     576
Arg Thr Gly Val Gly Arg Ile Leu Gln Val Met Val Trp Asn Ile Gly
            180                 185                 190
```

```
ctg gcg tac ggg gcg att tcg ctg gcg gac ctg gcc tgg cag cgt tac      624
Leu Ala Tyr Gly Ala Ile Ser Leu Ala Asp Leu Ala Trp Gln Arg Tyr
        195                 200                 205

cag tat cgc aaa ggc ttg cgg atg agc aag gac gaa gtg aag cag gag      672
Gln Tyr Arg Lys Gly Leu Arg Met Ser Lys Asp Glu Val Lys Gln Glu
    210                 215                 220

tac aag gag atg gaa ggc gat ccc cat atc aag cag caa cgc aag cac      720
Tyr Lys Glu Met Glu Gly Asp Pro His Ile Lys Gln Gln Arg Lys His
225                 230                 235                 240

ctg cac cag gag ctg atc atg cat ggc gcg gcg gcc cag gtt cgc cgg      768
Leu His Gln Glu Leu Ile Met His Gly Ala Ala Ala Gln Val Arg Arg
                245                 250                 255

gcg acg gtg ctg gtg acc aat ccg aca cac ctg gcc gtg gcc ctg tac      816
Ala Thr Val Leu Val Thr Asn Pro Thr His Leu Ala Val Ala Leu Tyr
            260                 265                 270

tac gcg gcg ggc gag acg ccc ttg ccg cgc gtg ctg gcc atg ggg cag      864
Tyr Ala Ala Gly Glu Thr Pro Leu Pro Arg Val Leu Ala Met Gly Gln
        275                 280                 285

gga gcc gtg gcc gct ctc atg gtc gag gcc gcg cgc gat gcc ggc gtg      912
Gly Ala Val Ala Ala Leu Met Val Glu Ala Ala Arg Asp Ala Gly Val
    290                 295                 300

ccg gtc atg cag aac gtc gcg ctg gcc cgc gcc ttg cac gac cag gcg      960
Pro Val Met Gln Asn Val Ala Leu Ala Arg Ala Leu His Asp Gln Ala
305                 310                 315                 320

gag gtg gac caa tac att ccc ggc gag ttg gtg gag ccg gtg gcc gcg     1008
Glu Val Asp Gln Tyr Ile Pro Gly Glu Leu Val Glu Pro Val Ala Ala
                325                 330                 335

gtg ttg cgg gcg gtg cgc cag gca ctc aag gag cag aca tga              1050
Val Leu Arg Ala Val Arg Gln Ala Leu Lys Glu Gln Thr  *
            340                 345
```

<210> 36
<211> 349
<212> PRT
<213> Bordetella pertussis

<400> 36

```
Met Ser Gly Glu Lys Thr Glu Arg Pro Thr Pro Lys Arg Leu Arg Asp
 1               5                  10                  15
Ser Arg Glu Lys Gly Glu Val Ala His Ser Arg Asp Phe Thr Gln Thr
            20                  25                  30
Ala Leu Ile Cys Ala Leu Phe Gly His Phe Leu Ile Asn Ala Pro Ser
        35                  40                  45
Ile Leu Ala Ser Leu Arg Ala Leu Ile Leu Ala Pro Ala Ala Phe Ala
    50                  55                  60
Asp Gln Gly Phe Ala Val Ala Leu Gly Pro Val Leu Thr Glu Ile Leu
65                  70                  75                  80
Asp Gln Ala Val Arg Val Leu Ala Pro Leu Ile Leu Ile Val Leu Gly
                85                  90                  95
```

```
Val Gly Met Phe Ala Glu Phe Leu Gln Val Gly Val Val Leu Ala Phe
                100                 105                 110
Arg Lys Leu Lys Pro Ser Ala Glu Lys Leu Asn Pro Ala Gly Asn Leu
            115                 120                 125
Lys Asn Ile Phe Ser Ala Arg Asn Leu Met Glu Phe Ile Lys Ser Val
    130                 135                 140
Cys Lys Ile Leu Phe Leu Ala Val Leu Val Thr Leu Val Ile Arg Asp
145                 150                 155                 160
Ser Leu Gln Pro Leu Met ala Val Pro His Ser Gly Leu Asp Gly Leu
                165                 170                 175
Arg Thr Gly Val Gly Arg Ile Leu Gln Val Met Val Trp Asn Ile Gly
            180                 185                 190
Leu Ala Tyr Gly Ala Ile Ser Leu Ala Asp Leu Ala Trp Gln Arg Tyr
        195                 200                 205
Gln Tyr Arg Lys Gly Leu Arg Met Ser Lys Asp Glu Val Lys Gln Glu
    210                 215                 220
Tyr Lys Glu Met Glu Gly Asp Pro His Ile Lys Gln Gln Arg Lys His
225                 230                 235                 240
Leu His Gln Glu Leu Ile Met His Gly Ala Ala Ala Gln Val Arg Arg
                245                 250                 255
Ala Thr Val Leu Val Thr Asn Pro Thr His Leu Ala Val Ala Leu Tyr
            260                 265                 270
Tyr Ala Ala Gly Glu Thr Pro Leu Pro Arg Val Leu Ala Met Gly Gln
        275                 280                 285
Gly Ala Val Ala Ala Leu Met Val Glu Ala Ala Arg Asp Ala Gly Val
    290                 295                 300
Pro Val Met Gln Asn Val Ala Leu Ala Arg Ala Leu His Asp Gln Ala
305                 310                 315                 320
Glu Val Asp Gln Tyr Ile Pro Gly Glu Leu Val Glu Pro Val Ala Ala
                325                 330                 335
Val Leu Arg Ala Val Arg Gln Ala Leu Lys Glu Gln Thr
            340                 345
```

<210> 37
<211> 399
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(399)

<400> 37

```
atg aca gca acc att cat ccc gat att gcc gat tat gcg cga cgc cat        48
Met Thr Ala Thr Ile His Pro Asp Ile Ala Asp Tyr Ala Arg Arg His
 1               5                  10                  15

ggc ctc gaa ccc tcg gtc gac gcc gat ggc ggg ctt gcc gtc cgg atc        96
Gly Leu Glu Pro Ser Val Asp Ala Asp Gly Gly Leu Ala Val Arg Ile
            20                  25                  30

gac gga cgg cat cgc gtc agg ttg atc ccc gcc gaa gac ggc atg ctg       144
Asp Gly Arg His Arg Val Arg Leu Ile Pro Ala Glu Asp Gly Met Leu
        35                  40                  45

gtg ttg cgg gcg cgg ctg gcc gag ctg ccc gat ggg tgg cag gcg cgc       192
Val Leu Arg Ala Arg Leu Ala Glu Leu Pro Asp Gly Trp Gln Ala Arg
    50                  55                  60

gcg gcg cag ttg cgc cgg gcg ggc ctg ctg gcc agc gcc atg gcc cct       240
```

```
Ala Ala Gln Leu Arg Arg Ala Gly Leu Leu Ala Ser Ala Met ala Pro
 65                  70                  75                  80

gcg acc gat gcg tac tgc ggc ata gac cag ggc gaa acc gcg ttg tat      288
Ala Thr Asp Ala Tyr Cys Gly Ile Asp Gln Gly Glu Thr Ala Leu Tyr
                85                  90                  95

ctg cac cag cgc gtc gca ccg gcc ggc agt gcg ctg gcg gtg gac gag      336
Leu His Gln Arg Val Ala Pro Ala Gly Ser Ala Leu Ala Val Asp Glu
            100                 105                 110

gcg gtg ggc gag ttc gtc aat gcc ttg gcc act tgg aaa agg gcg atg      384
Ala Val Gly Glu Phe Val Asn Ala Leu Ala Thr Trp Lys Arg Ala Met
        115                 120                 125

gcg caa tgg caa tag                                                 399
Ala Gln Trp Gln  *
    130
```

<210> 38
<211> 132
<212> PRT
<213> Bordetella pertussis

<400> 38

```
Met Thr Ala Thr Ile His Pro Asp Ile Ala Asp Tyr Ala Arg Arg His
 1               5                  10                  15
Gly Leu Glu Pro Ser Val Asp Ala Asp Gly Gly Leu Ala Val Arg Ile
            20                  25                  30
Asp Gly Arg His Arg Val Arg Leu Ile Pro Ala Glu Asp Gly Met Leu
        35                  40                  45
Val Leu Arg Ala Arg Leu Ala Glu Leu Pro Asp Gly Trp Gln Ala Arg
    50                  55                  60
Ala Ala Gln Leu Arg Arg Ala Gly Leu Leu Ala Ser Ala Met ala Pro
65                  70                  75                  80
Ala Thr Asp Ala Tyr Cys Gly Ile Asp Gln Gly Glu Thr Ala Leu Tyr
                85                  90                  95
Leu His Gln Arg Val Ala Pro Ala Gly Ser Ala Leu Ala Val Asp Glu
            100                 105                 110
Ala Val Gly Glu Phe Val Asn Ala Leu Ala Thr Trp Lys Arg Ala Met
        115                 120                 125
Ala Gln Trp Gln
    130
```

<210> 39
<211> 603
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(603)

<400> 39

```
atg gtt tct ccc ccg tca tcc ggt ctt ccc gct tct ctc gaa aaa ccg       48
Met Val Ser Pro Pro Ser Ser Gly Leu Pro Ala Ser Leu Glu Lys Pro
 1               5                  10                  15

gac aac gca tat ccc gat atc gcc acc gag cga tcg gac cag cag ttg       96
```

```
Asp Asn Ala Tyr Pro Asp Ile Ala Thr Glu Arg Ser Asp Gln Gln Leu
            20                  25                  30

ctg agc agc ctg gta gcc gaa cat gcc ggc cga tta cag aga ttc atc      144
Leu Ser Ser Leu Val Ala Glu His Ala Gly Arg Leu Gln Arg Phe Ile
        35                  40                  45

gcc aag cac atc ggc cac agc agc gac gtc gag gac ctt gcg cag cag      192
Ala Lys His Ile Gly His Ser Ser Asp Val Glu Asp Leu Ala Gln Gln
    50                  55                  60

gct ttc gcc gag gcg gcg cgc gcg tat caa tcg ttc cgt ggc gac tcc      240
Ala Phe Ala Glu Ala Ala Arg Ala Tyr Gln Ser Phe Arg Gly Asp Ser
65                  70                  75                  80

cag ctt tcc acc tgg ctg tac ggc atc gcg ctc aat ctg gtc cgc aat      288
Gln Leu Ser Thr Trp Leu Tyr Gly Ile Ala Leu Asn Leu Val Arg Asn
                85                  90                  95

cac ttg tcg cgt gcg cca gag cgc cgt tat gaa ttc acc tcc gac gcc      336
His Leu Ser Arg Ala Pro Glu Arg Arg Tyr Glu Phe Thr Ser Asp Ala
            100                 105                 110

agc ctg ggt gtc atg cca tgc agt gcg ccc aac ccc gaa gcc gtg acc      384
Ser Leu Gly Val Met Pro Cys Ser Ala Pro Asn Pro Glu Ala Val Thr
        115                 120                 125

gag cag cgt caa cgc atg cgc ttg cta cgc gaa gcg ctg gag cag ctc      432
Glu Gln Arg Gln Arg Met Arg Leu Leu Arg Glu Ala Leu Glu Gln Leu
    130                 135                 140

ccc gaa agc atg cgc gac gtg atc ctc atg gtc ggc gtg gaa gaa ctc      480
Pro Glu Ser Met Arg Asp Val Ile Leu Met Val Gly Val Glu Glu Leu
145                 150                 155                 160

tcc tat gaa gag gct gcc gca ctg ctt tcg gtt cct gta gga acc att      528
Ser Tyr Glu Glu Ala Ala Ala Leu Leu Ser Val Pro Val Gly Thr Ile
                165                 170                 175

cgc agc cga ctt tcc cgc gcc cgc tgt gcc ttg cgc gaa gcg ctg cgc      576
Arg Ser Arg Leu Ser Arg Ala Arg Cys Ala Leu Arg Glu Ala Leu Arg
            180                 185                 190

gaa cga ggc tac gac agc gtg ccg tag                                  603
Glu Arg Gly Tyr Asp Ser Val Pro  *
        195                 200


      <210> 40
      <211> 200
      <212> PRT
      <213> Bordetella pertussis

      <400> 40
Met Val Ser Pro Pro Ser Ser Gly Leu Pro Ala Ser Leu Glu Lys Pro
 1               5                  10                  15
Asp Asn Ala Tyr Pro Asp Ile Ala Thr Glu Arg Ser Asp Gln Gln Leu
            20                  25                  30
Leu Ser Ser Leu Val Ala Glu His Ala Gly Arg Leu Gln Arg Phe Ile
        35                  40                  45
```

```
Ala Lys His Ile Gly His Ser Ser Asp Val Glu Asp Leu Ala Gln Gln
    50                  55                  60
Ala Phe Ala Glu Ala Ala Arg Ala Tyr Gln Ser Phe Arg Gly Asp Ser
65                  70                  75                  80
Gln Leu Ser Thr Trp Leu Tyr Gly Ile Ala Leu Asn Leu Val Arg Asn
                85                  90                  95
His Leu Ser Arg Ala Pro Glu Arg Arg Tyr Glu Phe Thr Ser Asp Ala
            100                 105                 110
Ser Leu Gly Val Met Pro Cys Ser Ala Pro Asn Pro Glu Ala Val Thr
        115                 120                 125
Glu Gln Arg Gln Arg Met Arg Leu Leu Arg Glu Ala Leu Glu Gln Leu
    130                 135                 140
Pro Glu Ser Met Arg Asp Val Ile Leu Met Val Gly Val Glu Glu Leu
145                 150                 155                 160
Ser Tyr Glu Glu Ala Ala Ala Leu Leu Ser Val Pro Val Gly Thr Ile
                165                 170                 175
Arg Ser Arg Leu Ser Arg Ala Arg Cys Ala Leu Arg Glu Ala Leu Arg
            180                 185                 190
Glu Arg Gly Tyr Asp Ser Val Pro
        195                 200
```

<210> 41
<211> 1098
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1098)

<400> 41

```
atg act cgt atc gat gcc gcc ccc aat ccc ttc cac gcc gcc atg cag      48
Met Thr Arg Ile Asp Ala Ala Pro Asn Pro Phe His Ala Ala Met Gln
 1               5                  10                  15

ggg cgc cac gac gcc tcg gcc aac acc tcc tcc ggc tgg ctg caa ggc      96
Gly Arg His Asp Ala Ser Ala Asn Thr Ser Ser Gly Trp Leu Gln Gly
            20                  25                  30

cag cgc atc gca ccg gcg ccc acc ggc ata tcg ctg gcg gac gcg gcc     144
Gln Arg Ile Ala Pro Ala Pro Thr Gly Ile Ser Leu Ala Asp Ala Ala
        35                  40                  45

gag gag ctc agc ctg cac atg gcg cag gct gcc gag gaa aag cat cac     192
Glu Glu Leu Ser Leu His Met ala Gln Ala Ala Glu Glu Lys His His
    50                  55                  60

tcc gaa cgc aag gtc acg gcc gaa cgt ccg atg ctc tgg ctg gac gcg     240
Ser Glu Arg Lys Val Thr Ala Glu Arg Pro Met Leu Trp Leu Asp Ala
65                  70                  75                  80

gcg cag ctt gcg gaa ctg ttt tcc cac acc cac gac ccc gac gcg cag     288
Ala Gln Leu Ala Glu Leu Phe Ser His Thr His Asp Pro Asp Ala Gln
                85                  90                  95

gca aaa ctg gaa gcc ctg acc gcc gag ctg ctg cgc ggc cgg ggc gcc     336
Ala Lys Leu Glu Ala Leu Thr Ala Glu Leu Leu Arg Gly Arg Gly Ala
            100                 105                 110

ccc atg cag ctg gcc gcg caa gcg ttt ccc ggt gtc acg cag caa tac     384
```

```
Pro Met Gln Leu Ala Ala Gln Ala Phe Pro Gly Val Thr Gln Gln Tyr
        115                 120                 125

ctc gcg ctg cag cac gcg ctg cag cgc ggc gag cac gag gac gcc gcg      432
Leu Ala Leu Gln His Ala Leu Gln Arg Gly Glu His Glu Asp Ala Ala
    130                 135                 140

ccg cac gcg ctc gaa gcc ctg cgc gat gca ttg gcc gac ctg gag ctc      480
Pro His Ala Leu Glu Ala Leu Arg Asp Ala Leu Ala Asp Leu Glu Leu
145                 150                 155                 160

gcc cat ggc ccc gaa atc cgc gcc ggc atc aac acc ctg ccc acg gcc      528
Ala His Gly Pro Glu Ile Arg Ala Gly Ile Asn Thr Leu Pro Thr Ala
                165                 170                 175

ggc gca ttc gcg cgt tcc gct gac gag ctg gcc ggc ttc cag cac gcg      576
Gly Ala Phe Ala Arg Ser Ala Asp Glu Leu Ala Gly Phe Gln His Ala
            180                 185                 190

tac cgc gac atc gcc ctg ggc cag ctg tcg ttg gcg cgc acg ctg gac      624
Tyr Arg Asp Ile Ala Leu Gly Gln Leu Ser Leu Ala Arg Thr Leu Asp
        195                 200                 205

ctg gtg ctg gaa cgc tat ggg aac gac gac atc cac ggc gcg ctg ggc      672
Leu Val Leu Glu Arg Tyr Gly Asn Asp Asp Ile His Gly Ala Leu Gly
    210                 215                 220

gcg ctg att cag gcg ctg gga cac gac ctg gcc gcg gcg aca ccg tcg      720
Ala Leu Ile Gln Ala Leu Gly His Asp Leu Ala Ala Ala Thr Pro Ser
225                 230                 235                 240

acg gac ggc gtc agg ctg caa gtg ttg gcg agc gat ctc tat caa gtc      768
Thr Asp Gly Val Arg Leu Gln Val Leu Ala Ser Asp Leu Tyr Gln Val
                245                 250                 255

gag gtg gcc gcc acg gta ctg gag gaa tgc aat gcc ctg aaa caa cgg      816
Glu Val Ala Ala Thr Val Leu Glu Glu Cys Asn Ala Leu Lys Gln Arg
            260                 265                 270

ttg ggc aac gca ggc tcg cag gag tgt gcg gac gcc cag ggc ctg atg      864
Leu Gly Asn Ala Gly Ser Gln Glu Cys Ala Asp Ala Gln Gly Leu Met
        275                 280                 285

cgc gat ctt gtg gga atc agc gag gac aaa tgg att gcg ccc gcg cgc      912
Arg Asp Leu Val Gly Ile Ser Glu Asp Lys Trp Ile Ala Pro Ala Arg
    290                 295                 300

ttc gag aag ctg gcc gag cgc cac ggt gcg aat gcc ctc tcc gag cgc      960
Phe Glu Lys Leu Ala Glu Arg His Gly Ala Asn Ala Leu Ser Glu Arg
305                 310                 315                 320

atc gca ttc ctc ggt ggc gta cgc cag att ctc aaa gac ctg ccc acg     1008
Ile Ala Phe Leu Gly Gly Val Arg Gln Ile Leu Lys Asp Leu Pro Thr
                325                 330                 335

cag atc tac gcc gac atg gac gtg cgc gcc acc gtc ctg gcg gcc gcg     1056
Gln Ile Tyr Ala Asp Met Asp Val Arg Ala Thr Val Leu Ala Ala Ala
            340                 345                 350

cag gac gcg ctg gac aac gcg ata gca atg gag aac gca tga             1098
```

```
Gln Asp Ala Leu Asp Asn Ala Ile Ala Met Glu Asn Ala  *
        355                 360                 365


      <210> 42
      <211> 365
      <212> PRT
      <213> Bordetella pertussis

      <400> 42
Met Thr Arg Ile Asp Ala Ala Pro Asn Pro Phe His Ala Ala Met Gln
 1               5                  10                  15
Gly Arg His Asp Ala Ser Ala Asn Thr Ser Ser Gly Trp Leu Gln Gly
            20                  25                  30
Gln Arg Ile Ala Pro Ala Pro Thr Gly Ile Ser Leu Ala Asp Ala Ala
        35                  40                  45
Glu Glu Leu Ser Leu His Met ala Gln Ala Ala Glu Glu Lys His His
    50                  55                  60
Ser Glu Arg Lys Val Thr Ala Glu Arg Pro Met Leu Trp Leu Asp Ala
65                  70                  75                  80
Ala Gln Leu Ala Glu Leu Phe Ser His Thr His Asp Pro Asp Ala Gln
                85                  90                  95
Ala Lys Leu Glu Ala Leu Thr Ala Glu Leu Leu Arg Gly Arg Gly Ala
            100                 105                 110
Pro Met Gln Leu Ala Ala Gln Ala Phe Pro Gly Val Thr Gln Gln Tyr
        115                 120                 125
Leu Ala Leu Gln His Ala Leu Gln Arg Gly Glu His Glu Asp Ala Ala
    130                 135                 140
Pro His Ala Leu Glu Ala Leu Arg Asp Ala Leu Ala Asp Leu Glu Leu
145                 150                 155                 160
Ala His Gly Pro Glu Ile Arg Ala Gly Ile Asn Thr Leu Pro Thr Ala
                165                 170                 175
Gly Ala Phe Ala Arg Ser Ala Asp Glu Leu Ala Gly Phe Gln His Ala
            180                 185                 190
Tyr Arg Asp Ile Ala Leu Gly Gln Leu Ser Leu Ala Arg Thr Leu Asp
        195                 200                 205
Leu Val Leu Glu Arg Tyr Gly Asn Asp Asp Ile His Gly Ala Leu Gly
    210                 215                 220
Ala Leu Ile Gln Ala Leu Gly His Asp Leu Ala Ala Ala Thr Pro Ser
225                 230                 235                 240
Thr Asp Gly Val Arg Leu Gln Val Leu Ala Ser Asp Leu Tyr Gln Val
                245                 250                 255
Glu Val Ala Ala Thr Val Leu Glu Glu Cys Asn Ala Leu Lys Gln Arg
            260                 265                 270
Leu Gly Asn Ala Gly Ser Gln Glu Cys Ala Asp Ala Gln Gly Leu Met
        275                 280                 285
Arg Asp Leu Val Gly Ile Ser Glu Asp Lys Trp Ile Ala Pro Ala Arg
    290                 295                 300
Phe Glu Lys Leu Ala Glu Arg His Gly Ala Asn Ala Leu Ser Glu Arg
305                 310                 315                 320
Ile Ala Phe Leu Gly Gly Val Arg Gln Ile Leu Lys Asp Leu Pro Thr
                325                 330                 335
Gln Ile Tyr Ala Asp Met Asp Val Arg Ala Thr Val Leu Ala Ala Ala
            340                 345                 350
Gln Asp Ala Leu Asp Asn Ala Ile Ala Met Glu Asn Ala
        355                 360                 365

      <210> 43
      <211> 588
      <212> DNA
```

<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(588)

<400> 43

```
atg ccc ttg gcg ggg tcc tgc ggc cgg acg tcc gaa aac cgg gat att        48
Met Pro Leu Ala Gly Ser Cys Gly Arg Thr Ser Glu Asn Arg Asp Ile
 1               5                  10                  15

ccg aga agc aat cgg ctg gcg cct tcc aga ctg tgc gca tac cat tgc        96
Pro Arg Ser Asn Arg Leu Ala Pro Ser Arg Leu Cys Ala Tyr His Cys
            20                  25                  30

cct ctt ttg cca cgc att tcg agc gta tgg ttt ccc ttg cgc ccg acg       144
Pro Leu Leu Pro Arg Ile Ser Ser Val Trp Phe Pro Leu Arg Pro Thr
        35                  40                  45

cag gct cgc ctc gcc atg acc gac acg gca tac cac caa ctc atc gcc       192
Gln Ala Arg Leu Ala Met Thr Asp Thr Ala Tyr His Gln Leu Ile Ala
    50                  55                  60

gat ttc ggc cgc ctc atc ggc atc gac tcg ctc aac ccc ggt gcc ggc       240
Asp Phe Gly Arg Leu Ile Gly Ile Asp Ser Leu Asn Pro Gly Ala Gly
65                  70                  75                  80

ggc ctg tgt cag ttg att ttc gaa ccg tgc gca ccg gtc ttc atc gca       288
Gly Leu Cys Gln Leu Ile Phe Glu Pro Cys Ala Pro Val Phe Ile Ala
                85                  90                  95

ccg gtg cac gcc cgg acg gaa atc atg att tcc tgc gtg ctg ggc acg       336
Pro Val His Ala Arg Thr Glu Ile Met Ile Ser Cys Val Leu Gly Thr
            100                 105                 110

gcg gac gcg gcc aac ccg gca agc atg gcc cga gcc aac ttc atg cag       384
Ala Asp Ala Ala Asn Pro Ala Ser Met ala Arg Ala Asn Phe Met Gln
        115                 120                 125

gcc ggc agc ggc gtc gtg gcc tgc atc ggc ggc gat ggg ttg ttc tat       432
Ala Gly Ser Gly Val Val Ala Cys Ile Gly Gly Asp Gly Leu Phe Tyr
    130                 135                 140

ctg cag cag gcc ata ccc ctg tcg cgc gcc acg ccc gca atc ctg ctc       480
Leu Gln Gln Ala Ile Pro Leu Ser Arg Ala Thr Pro Ala Ile Leu Leu
145                 150                 155                 160

gat cac tgt gag cgt ctg ctg cag gaa gcc tcg cgc tgg cgc gtc ggc       528
Asp His Cys Glu Arg Leu Leu Gln Glu Ala Ser Arg Trp Arg Val Gly
                165                 170                 175

gac cac gac ggc tgc gcc acc tcg gcc ccg aat atc gcc gcg ctg acg       576
Asp His Asp Gly Cys Ala Thr Ser Ala Pro Asn Ile Ala Ala Leu Thr
            180                 185                 190

cgc ggc gtc tag                                                      588
Arg Gly Val  *
        195
```

<210> 44
<211> 195
<212> PRT
<213> Bordetella pertussis

<400> 44

```
Met Pro Leu Ala Gly Ser Cys Gly Arg Thr Ser Glu Asn Arg Asp Ile
 1               5                  10                  15
Pro Arg Ser Asn Arg Leu Ala Pro Ser Arg Leu Cys Ala Tyr His Cys
            20                  25                  30
Pro Leu Leu Pro Arg Ile Ser Ser Val Trp Phe Pro Leu Arg Pro Thr
        35                  40                  45
Gln Ala Arg Leu Ala Met Thr Asp Thr Ala Tyr His Gln Leu Ile Ala
    50                  55                  60
Asp Phe Gly Arg Leu Ile Gly Ile Asp Ser Leu Asn Pro Gly Ala Gly
65                  70                  75                  80
Gly Leu Cys Gln Leu Ile Phe Glu Pro Cys Ala Pro Val Phe Ile Ala
                85                  90                  95
Pro Val His Ala Arg Thr Glu Ile Met Ile Ser Cys Val Leu Gly Thr
            100                 105                 110
Ala Asp Ala Ala Asn Pro Ala Ser Met ala Arg Ala Asn Phe Met Gln
        115                 120                 125
Ala Gly Ser Gly Val Val Ala Cys Ile Gly Gly Asp Gly Leu Phe Tyr
    130                 135                 140
Leu Gln Gln Ala Ile Pro Leu Ser Arg Ala Thr Pro Ala Ile Leu Leu
145                 150                 155                 160
Asp His Cys Glu Arg Leu Leu Gln Glu Ala Ser Arg Trp Arg Val Gly
                165                 170                 175
Asp His Asp Gly Cys Ala Thr Ser Ala Pro Asn Ile Ala Ala Leu Thr
            180                 185                 190
Arg Gly Val
        195
```

<210> 45
<211> 369
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(369)

<400> 45

```
atg atg ccg cat acc cta ccc tcg ccc agc ctt cag gta cgc gaa ctt      48
Met Met Pro His Thr Leu Pro Ser Pro Ser Leu Gln Val Arg Glu Leu
 1               5                  10                  15

ctg caa ttg ctt gcg cac cac tac cag ttg cag cgc caa tgg agc aag      96
Leu Gln Leu Leu Ala His His Tyr Gln Leu Gln Arg Gln Trp Ser Lys
            20                  25                  30

acg gtt gcc ctg ctg gcg gcc ctg gat gcc ctg gac gcg atc gac agc     144
Thr Val Ala Leu Leu Ala Ala Leu Asp Ala Leu Asp Ala Ile Asp Ser
        35                  40                  45

cag tcc ctg ctg gcc ctg gcg ctg ggc tat ctg cac cag ggc gaa ccg     192
Gln Ser Leu Leu Ala Leu Ala Leu Gly Tyr Leu His Gln Gly Glu Pro
    50                  55                  60

cgc atg gcc ttg gtc acg ctg gac aag cgc gca ctg cgc gcc aca ccc     240
```

Arg Met *ala* Leu Val Thr Leu Asp Lys Arg Ala Leu Arg Ala Thr Pro
65 70 75 80

gat gcc gcc ggc cat ctg gtc cgg gcg cag gcc atg cag gcg ctg aac 288
Asp Ala Ala Gly His Leu Val Arg Ala Gln Ala Met Gln Ala Leu Asn
85 90 95

cga ccc gac gac gcc cgc cag gcc atg cgt gac tat atg gcg cta cgg 336
Arg Pro Asp Asp Ala Arg Gln Ala Met Arg Asp Tyr Met *ala* Leu Arg
100 105 110

gcg gcg tct tgc gcg agc gcg acc ccg cca tga 369
Ala Ala Ser Cys Ala Ser Ala Thr Pro Pro *
115 120

<210> 46
<211> 122
<212> PRT
<213> Bordetella pertussis

<400> 46
Met Met Pro His Thr Leu Pro Ser Pro Ser Leu Gln Val Arg Glu Leu
1 5 10 15
Leu Gln Leu Leu Ala His His Tyr Gln Leu Gln Arg Gln Trp Ser Lys
20 25 30
Thr Val Ala Leu Leu Ala Ala Leu Asp Ala Leu Asp Ala Ile Asp Ser
35 40 45
Gln Ser Leu Leu Ala Leu Ala Leu Gly Tyr Leu His Gln Gly Glu Pro
50 55 60
Arg Met *ala* Leu Val Thr Leu Asp Lys Arg Ala Leu Arg Ala Thr Pro
65 70 75 80
Asp Ala Ala Gly His Leu Val Arg Ala Gln Ala Met Gln Ala Leu Asn
85 90 95
Arg Pro Asp Asp Ala Arg Gln Ala Met Arg Asp Tyr Met *ala* Leu Arg
100 105 110
Ala Ala Ser Cys Ala Ser Ala Thr Pro Pro
115 120

<210> 47
<211> 411
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(411)

<400> 47
atg tcc agc gcc gta ccc ggc atg cat gcc atg cac ctc ggc ctg gag 48
Met Ser Ser Ala Val Pro Gly Met His Ala Met His Leu Gly Leu Glu
1 5 10 15

cgc ggc gtc gac cac atc gtg cgc ggt ccc cgc tgc gag ccc gcc ccc 96
Arg Gly Val Asp His Ile Val Arg Gly Pro Arg Cys Glu Pro Ala Pro
20 25 30

acc ctg cca ccc gag cgc tgg ctc gaa ccg ccc gcc acc ggc gcg gtc 144
Thr Leu Pro Pro Glu Arg Trp Leu Glu Pro Pro Ala Thr Gly Ala Val
35 40 45

```
gat cat ctg aaa gcc ctg ctc gta cgc ccg gac ctg agc gcg atg ctc      192
Asp His Leu Lys Ala Leu Leu Val Arg Pro Asp Leu Ser Ala Met Leu
     50                  55                  60

gac gag tcg gcg cgg cct cgc ctg acg gat ggc gca tta ttc cag ccc      240
Asp Glu Ser Ala Arg Pro Arg Leu Thr Asp Gly Ala Leu Phe Gln Pro
 65                  70                  75                  80

gcg cag ttc gag cgc gcc ctg gcc cag gcg cgc gac gaa ctg tcc cgg      288
Ala Gln Phe Glu Arg Ala Leu Ala Gln Ala Arg Asp Glu Leu Ser Arg
                 85                  90                  95

gcc atg gaa ctg cat gcc ggc aac acc gcg cca gcc tta agc cgc gcc      336
Ala Met Glu Leu His Ala Gly Asn Thr Ala Pro Ala Leu Ser Arg Ala
            100                 105                 110

ttg cac gta ctc aac gag gcc gga aag ctg cgc gac ctg gct gcc atg      384
Leu His Val Leu Asn Glu Ala Gly Lys Leu Arg Asp Leu Ala Ala Met
        115                 120                 125

tat cgc agc gcg ctc tac cag gga tga                                  411
Tyr Arg Ser Ala Leu Tyr Gln Gly  *
    130                 135
```

<210> 48
<211> 136
<212> PRT
<213> Bordetella pertussis

<400> 48

```
Met Ser Ser Ala Val Pro Gly Met His Ala Met His Leu Gly Leu Glu
 1               5                  10                  15
Arg Gly Val Asp His Ile Val Arg Gly Pro Arg Cys Glu Pro Ala Pro
            20                  25                  30
Thr Leu Pro Pro Glu Arg Trp Leu Glu Pro Pro Ala Thr Gly Ala Val
        35                  40                  45
Asp His Leu Lys Ala Leu Leu Val Arg Pro Asp Leu Ser Ala Met Leu
    50                  55                  60
Asp Glu Ser Ala Arg Pro Arg Leu Thr Asp Gly Ala Leu Phe Gln Pro
65                  70                  75                  80
Ala Gln Phe Glu Arg Ala Leu Ala Gln Ala Arg Asp Glu Leu Ser Arg
                85                  90                  95
Ala Met Glu Leu His Ala Gly Asn Thr Ala Pro Ala Leu Ser Arg Ala
            100                 105                 110
Leu His Val Leu Asn Glu Ala Gly Lys Leu Arg Asp Leu Ala Ala Met
        115                 120                 125
Tyr Arg Ser Ala Leu Tyr Gln Gly
    130                 135
```

<210> 49
<211> 378
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(378)

<400> 49

```
atg aat act gcc gat agg gcg ctg cat cag ttc ggc cag gat atc ggc        48
Met Asn Thr Ala Asp Arg Ala Leu His Gln Phe Gly Gln Asp Ile Gly
 1               5                  10                  15

atc gag ggc ctg gca ttc ggg ccg tcc gga tcg gcg tcg ctg gcg ctg        96
Ile Glu Gly Leu Ala Phe Gly Pro Ser Gly Ser Ala Ser Leu Ala Leu
            20                  25                  30

tcc aac ggg cgc cgc ctg ggc gtc gaa tgc gtc gcc ggc gcg gcc ctg       144
Ser Asn Gly Arg Arg Leu Gly Val Glu Cys Val Ala Gly Ala Ala Leu
        35                  40                  45

gtc cac ctg gcc cag cgg gtc gag cgc gac gcc gcc tcc gtg ttg ctg       192
Val His Leu Ala Gln Arg Val Glu Arg Asp Ala Ala Ser Val Leu Leu
    50                  55                  60

gcg gca tgg aaa cgg gcc cat ggg cag cgc gga agc gcc gca tcc atc       240
Ala Ala Trp Lys Arg Ala His Gly Gln Arg Gly Ser Ala Ala Ser Ile
65                  70                  75                  80

cag acg tca ctc tgg tcg gag ggc agc gag gac tgg atc gtc gcg cag       288
Gln Thr Ser Leu Trp Ser Glu Gly Ser Glu Asp Trp Ile Val Ala Gln
                85                  90                  95

aca cga ctg ccc gaa cgc tcg ctc gac gca gcg gcg ttg cgc ctg gcg       336
Thr Arg Leu Pro Glu Arg Ser Leu Asp Ala Ala Ala Leu Arg Leu Ala
            100                 105                 110

gtg ctg ggc ctg acg aac tgg ctc gac cgc ctg gag gcg tga               378
Val Leu Gly Leu Thr Asn Trp Leu Asp Arg Leu Glu Ala  *
        115                 120                 125
```

<210> 50
<211> 125
<212> PRT
<213> Bordetella pertussis

<400> 50

```
Met Asn Thr Ala Asp Arg Ala Leu His Gln Phe Gly Gln Asp Ile Gly
 1               5                  10                  15
Ile Glu Gly Leu Ala Phe Gly Pro Ser Gly Ser Ala Ser Leu Ala Leu
            20                  25                  30
Ser Asn Gly Arg Arg Leu Gly Val Glu Cys Val Ala Gly Ala Ala Leu
        35                  40                  45
Val His Leu Ala Gln Arg Val Glu Arg Asp Ala Ala Ser Val Leu Leu
    50                  55                  60
Ala Ala Trp Lys Arg Ala His Gly Gln Arg Gly Ser Ala Ala Ser Ile
65                  70                  75                  80
Gln Thr Ser Leu Trp Ser Glu Gly Ser Glu Asp Trp Ile Val Ala Gln
                85                  90                  95
Thr Arg Leu Pro Glu Arg Ser Leu Asp Ala Ala Ala Leu Arg Leu Ala
            100                 105                 110
Val Leu Gly Leu Thr Asn Trp Leu Asp Arg Leu Glu Ala
        115                 120                 125
```

<210> 51
<211> 783
<212> DNA

<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(783)

<400> 51

```
atg ggg agt cct cgg aga agg aac cat ttg cct act ggt gca gtg agt        48
Met Gly Ser Pro Arg Arg Arg Asn His Leu Pro Thr Gly Ala Val Ser
 1               5                  10                  15

gtc gcg cgc gcg gtc atg gtt ccc gga aac ggg cgc gat att ggg caa        96
Val Ala Arg Ala Val Met Val Pro Gly Asn Gly Arg Asp Ile Gly Gln
            20                  25                  30

ttc gca gcc tgg aac ttg ccg cgg gcg cag ggt tac tca gca tgc gtc       144
Phe Ala Ala Trp Asn Leu Pro Arg Ala Gln Gly Tyr Ser Ala Cys Val
        35                  40                  45

ttt caa ctc gaa gga gct ctc atg agc att gat ctc gga gtt tca ctc       192
Phe Gln Leu Glu Gly Ala Leu Met Ser Ile Asp Leu Gly Val Ser Leu
    50                  55                  60

acg tcg cag gcc ggc ggc ctg caa ggc atc gac ctc aag agc atg gat       240
Thr Ser Gln Ala Gly Gly Leu Gln Gly Ile Asp Leu Lys Ser Met Asp
65                  70                  75                  80

atc cag act ctc atg gtg tat gtg cag ggt cgt cgc gcc gaa ctc ctc       288
Ile Gln Thr Leu Met Val Tyr Val Gln Gly Arg Arg Ala Glu Leu Leu
                85                  90                  95

acg gct caa atg cag acc cag gcc gaa gtg gtg cag aag gcc aat gaa       336
Thr Ala Gln Met Gln Thr Gln Ala Glu Val Val Gln Lys Ala Asn Glu
            100                 105                 110

cgc atg gcg cag ctc aac gag gtc ctg tcc gcg ctg tcc cgg gcc aag       384
Arg Met ala Gln Leu Asn Glu Val Leu Ser Ala Leu Ser Arg Ala Lys
        115                 120                 125

gcc gag ttt ccg ccc aat ccg aag ccg ggc gac acc atc ccg ggc tgg       432
Ala Glu Phe Pro Pro Asn Pro Lys Pro Gly Asp Thr Ile Pro Gly Trp
    130                 135                 140

gac agc cag aag atc agc cgg atc gag gtt cct ctc aat gat gcg ctg       480
Asp Ser Gln Lys Ile Ser Arg Ile Glu Val Pro Leu Asn Asp Ala Leu
145                 150                 155                 160

cgt gcc gcc ggc ctg acg ggc atg ttc gaa gcg cgc gat ggc cgg gtg       528
Arg Ala Ala Gly Leu Thr Gly Met Phe Glu Ala Arg Asp Gly Arg Val
                165                 170                 175

acc ggc ccc gac ggc cgg ggt acg cag gtc gtg aac ggc acg ggc gtc       576
Thr Gly Pro Asp Gly Arg Gly Thr Gln Val Val Asn Gly Thr Gly Val
            180                 185                 190

atg gcc ggt tcc acg acc tat aag gaa ctc gaa agt gcc tac acc acc       624
Met ala Gly Ser Thr Thr Tyr Lys Glu Leu Glu Ser Ala Tyr Thr Thr
        195                 200                 205

gta aag ggg atg ctg gat acg gcg tcc aat acg caa cag atg gac atg       672
```

```
Val Lys Gly Met Leu Asp Thr Ala Ser Asn Thr Gln Gln Met Asp Met
    210                 215                 220

atc agg ctg cag gcc gcc agc aac aag cgc aac gag gct ttc gag gtc      720
Ile Arg Leu Gln Ala Ala Ser Asn Lys Arg Asn Glu Ala Phe Glu Val
225                 230                 235                 240

atg acc aac acc gag aag cgg cgc agc gac ttg aac agc tcc atc acc      768
Met Thr Asn Thr Glu Lys Arg Arg Ser Asp Leu Asn Ser Ser Ile Thr
                245                 250                 255

agc aac atg cgc taa                                                  783
Ser Asn Met Arg  *
            260
```

<210> 52
<211> 260
<212> PRT
<213> Bordetella pertussis

<400> 52

```
Met Gly Ser Pro Arg Arg Arg Asn His Leu Pro Thr Gly Ala Val Ser
 1               5                  10                  15
Val Ala Arg Ala Val Met Val Pro Gly Asn Gly Arg Asp Ile Gly Gln
            20                  25                  30
Phe Ala Ala Trp Asn Leu Pro Arg Ala Gln Gly Tyr Ser Ala Cys Val
        35                  40                  45
Phe Gln Leu Glu Gly Ala Leu Met Ser Ile Asp Leu Gly Val Ser Leu
    50                  55                  60
Thr Ser Gln Ala Gly Gly Leu Gln Gly Ile Asp Leu Lys Ser Met Asp
65                  70                  75                  80
Ile Gln Thr Leu Met Val Tyr Val Gln Gly Arg Arg Ala Glu Leu Leu
                85                  90                  95
Thr Ala Gln Met Gln Thr Gln Ala Glu Val Val Gln Lys Ala Asn Glu
            100                 105                 110
Arg Met ala Gln Leu Asn Glu Val Leu Ser Ala Leu Ser Arg Ala Lys
        115                 120                 125
Ala Glu Phe Pro Pro Asn Pro Lys Pro Gly Asp Thr Ile Pro Gly Trp
    130                 135                 140
Asp Ser Gln Lys Ile Ser Arg Ile Glu Val Pro Leu Asn Asp Ala Leu
145                 150                 155                 160
Arg Ala Ala Gly Leu Thr Gly Met Phe Glu Ala Arg Asp Gly Arg Val
                165                 170                 175
Thr Gly Pro Asp Gly Arg Gly Thr Gln Val Val Asn Gly Thr Gly Val
            180                 185                 190
Met ala Gly Ser Thr Thr Tyr Lys Glu Leu Glu Ser Ala Tyr Thr Thr
        195                 200                 205
Val Lys Gly Met Leu Asp Thr Ala Ser Asn Thr Gln Gln Met Asp Met
    210                 215                 220
Ile Arg Leu Gln Ala Ala Ser Asn Lys Arg Asn Glu Ala Phe Glu Val
225                 230                 235                 240
Met Thr Asn Thr Glu Lys Arg Arg Ser Asp Leu Asn Ser Ser Ile Thr
                245                 250                 255
Ser Asn Met Arg
            260
```

<210> 53
<211> 276
<212> DNA

<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(276)

<400> 53

```
atg cag gag caa ggc atc caa tcc atc atg cgc gcc gcg gaa gag ctg        48
Met Gln Glu Gln Gly Ile Gln Ser Ile Met Arg Ala Ala Glu Glu Leu
 1               5                  10                  15

gtc gag cag acc cgc cag gcg ttg tac agc gtc gac gag atc tac gcc        96
Val Glu Gln Thr Arg Gln Ala Leu Tyr Ser Val Asp Glu Ile Tyr Ala
            20                  25                  30

cac gtt ggc gtc gac ccc gct cgc ctg cgc aat ctg gcg gtc gag cag       144
His Val Gly Val Asp Pro Ala Arg Leu Arg Asn Leu Ala Val Glu Gln
        35                  40                  45

gcc agg ata gag gcc gag gcc cag gcg gcg ttc cgt gat gac ctc gcg       192
Ala Arg Ile Glu Ala Glu Ala Gln Ala Ala Phe Arg Asp Asp Leu Ala
    50                  55                  60

gac atc gag cgc gag gcg gcg cgc gtc aag gcg gcc tgc acc gat gcg       240
Asp Ile Glu Arg Glu Ala Ala Arg Val Lys Ala Ala Cys Thr Asp Ala
65                  70                  75                  80

ccg cag gcc cgc agg gtg ctt cac aac cac gtc tga                      276
Pro Gln Ala Arg Arg Val Leu His Asn His Val  *
                85                  90
```

<210> 54
<211> 91
<212> PRT
<213> Bordetella pertussis

<400> 54

```
Met Gln Glu Gln Gly Ile Gln Ser Ile Met Arg Ala Ala Glu Glu Leu
 1               5                  10                  15
Val Glu Gln Thr Arg Gln Ala Leu Tyr Ser Val Asp Glu Ile Tyr Ala
            20                  25                  30
His Val Gly Val Asp Pro Ala Arg Leu Arg Asn Leu Ala Val Glu Gln
        35                  40                  45
Ala Arg Ile Glu Ala Glu Ala Gln Ala Ala Phe Arg Asp Asp Leu Ala
    50                  55                  60
Asp Ile Glu Arg Glu Ala Ala Arg Val Lys Ala Ala Cys Thr Asp Ala
65                  70                  75                  80
Pro Gln Ala Arg Arg Val Leu His Asn His Val
                85                  90
```

<210> 55
<211> 942
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(942)

<400> 55

```
atg tct gtt tct ccg act tcg ccc ggc tct ttc ggg gcc ggc cct gtc        48
Met Ser Val Ser Pro Thr Ser Pro Gly Ser Phe Gly Ala Gly Pro Val
 1               5                  10                  15

ttt gac tcc gaa ttg cag gcc ccg gcc ccg tcg gcg cag cgt cgc ggc        96
Phe Asp Ser Glu Leu Gln Ala Pro Ala Pro Ser Ala Gln Arg Arg Gly
            20                  25                  30

ggt gcg gcg cct gtg ccg ccg ccc gtc gat cgg cgc ggc gtc gag ccg       144
Gly Ala Ala Pro Val Pro Pro Pro Val Asp Arg Arg Gly Val Glu Pro
        35                  40                  45

gga gat ccc acg ctg ggc atg ctg ccc gcg cca gat ttg ctc gcg ggg       192
Gly Asp Pro Thr Leu Gly Met Leu Pro Ala Pro Asp Leu Leu Ala Gly
    50                  55                  60

ggc gcc gtc agc cgc acc cgc gcg gcg ctc gac gat ctg gac gcc gca       240
Gly Ala Val Ser Arg Thr Arg Ala Ala Leu Asp Asp Leu Asp Ala Ala
65                  70                  75                  80

cgg ctc ggt gaa gac atc tac gcc ttg atg gcg gtg ttg caa cag gcc       288
Arg Leu Gly Glu Asp Ile Tyr Ala Leu Met ala Val Leu Gln Gln Ala
                85                  90                  95

agt cag cag atg cgg gac gcc gcc cgt atc gct cgt gat gcc gag gct       336
Ser Gln Gln Met Arg Asp Ala Ala Arg Ile Ala Arg Asp Ala Glu Ala
            100                 105                 110

acg cgg caa acg cag gct ctc ggc gat gcg gcc agc cag atg cgc cag       384
Thr Arg Gln Thr Gln Ala Leu Gly Asp Ala Ala Ser Gln Met Arg Gln
        115                 120                 125

gcg gcg agc gag cgc atg gcc gga gcg atc gtg gcg ggc gcc atg cag       432
Ala Ala Ser Glu Arg Met ala Gly Ala Ile Val Ala Gly Ala Met Gln
    130                 135                 140

ata gcg ggt ggt ttc gtg cag ctg ggg gcg ggc ctg gca gcg ggt ttg       480
Ile Ala Gly Gly Phe Val Gln Leu Gly Ala Gly Leu Ala Ala Gly Leu
145                 150                 155                 160

cag gcc atg ggt ggc gct gct gcg caa gcc aag ggc gcc gca ttc tcc       528
Gln Ala Met Gly Gly Ala Ala Ala Gln Ala Lys Gly Ala Ala Phe Ser
                165                 170                 175

gag cag gcc tcg aca agc cgc aag gtg gcg gcc ggc ttg cac gat gcc       576
Glu Gln Ala Ser Thr Ser Arg Lys Val Ala Ala Gly Leu His Asp Ala
            180                 185                 190

ccc gag ctg cag gca acg gtg cag gcc cgc gca acc cag ctc gaa gcg       624
Pro Glu Leu Gln Ala Thr Val Gln Ala Arg Ala Thr Gln Leu Glu Ala
        195                 200                 205

caa gcg gcc tcg ttt ggt gcg gac gcg gct cgt tcg tcg gca aag tcg       672
Gln Ala Ala Ser Phe Gly Ala Asp Ala Ala Arg Ser Ser Ala Lys Ser
    210                 215                 220

cag cgc gta tcg agc gtt gcc cag gcc ggc gcc gca gcg gcc ggc ggt       720
Gln Arg Val Ser Ser Val Ala Gln Ala Gly Ala Ala Ala Ala Gly Gly
225                 230                 235                 240
```

```
atc ggc ggc ctg acc agc gcc gcc cag gaa cgc cgc gcc gcc gag cac      768
Ile Gly Gly Leu Thr Ser Ala Ala Gln Glu Arg Arg Ala Ala Glu His
                245                 250                 255

gag gcc agg cgc gcg gag ctg gac gtc gaa gcg aag gtg cat gaa acg      816
Glu Ala Arg Arg Ala Glu Leu Asp Val Glu Ala Lys Val His Glu Thr
            260                 265                 270

gcc tcg cgg cgg gcc gac gaa gcc atg cag cag atg ctc gac atc atc      864
Ala Ser Arg Arg Ala Asp Glu Ala Met Gln Gln Met Leu Asp Ile Ile
        275                 280                 285

cgc ggc atc agg gaa aag ctg gcc ggg atg gag cag tcc cgc agc gag      912
Arg Gly Ile Arg Glu Lys Leu Ala Gly Met Glu Gln Ser Arg Ser Glu
    290                 295                 300

acc gcc cgt agc gtg gcc cgc aat atc tga                              942
Thr Ala Arg Ser Val Ala Arg Asn Ile  *
305                 310
```

<210> 56
<211> 313
<212> PRT
<213> Bordetella pertussis

<400> 56

```
Met Ser Val Ser Pro Thr Ser Pro Gly Ser Phe Gly Ala Gly Pro Val
 1               5                  10                  15
Phe Asp Ser Glu Leu Gln Ala Pro Ala Pro Ser Ala Gln Arg Arg Gly
            20                  25                  30
Gly Ala Ala Pro Val Pro Pro Pro Val Asp Arg Arg Gly Val Glu Pro
        35                  40                  45
Gly Asp Pro Thr Leu Gly Met Leu Pro Ala Pro Asp Leu Leu Ala Gly
    50                  55                  60
Gly Ala Val Ser Arg Thr Arg Ala Ala Leu Asp Asp Leu Asp Ala Ala
65                  70                  75                  80
Arg Leu Gly Glu Asp Ile Tyr Ala Leu Met ala Val Leu Gln Gln Ala
                85                  90                  95
Ser Gln Gln Met Arg Asp Ala Ala Arg Ile Ala Arg Asp Ala Glu Ala
            100                 105                 110
Thr Arg Gln Thr Gln Ala Leu Gly Asp Ala Ala Ser Gln Met Arg Gln
        115                 120                 125
Ala Ala Ser Glu Arg Met ala Gly Ala Ile Val Ala Gly Ala Met Gln
    130                 135                 140
Ile Ala Gly Gly Phe Val Gln Leu Gly Ala Gly Leu Ala Ala Gly Leu
145                 150                 155                 160
Gln Ala Met Gly Gly Ala Ala Ala Gln Ala Lys Gly Ala Ala Phe Ser
                165                 170                 175
Glu Gln Ala Ser Thr Ser Arg Lys Val Ala Ala Gly Leu His Asp Ala
            180                 185                 190
Pro Glu Leu Gln Ala Thr Val Gln Ala Arg Ala Thr Gln Leu Glu Ala
        195                 200                 205
Gln Ala Ala Ser Phe Gly Ala Asp Ala Ala Arg Ser Ser Ala Lys Ser
    210                 215                 220
Gln Arg Val Ser Ser Val Ala Gln Ala Gly Ala Ala Ala Ala Gly Gly
225                 230                 235                 240
Ile Gly Gly Leu Thr Ser Ala Ala Gln Glu Arg Arg Ala Ala Glu His
                245                 250                 255
```

```
Glu Ala Arg Arg Ala Glu Leu Asp Val Glu Ala Lys Val His Glu Thr
            260                 265                 270
Ala Ser Arg Arg Ala Asp Glu Ala Met Gln Gln Met Leu Asp Ile Ile
        275                 280                 285
Arg Gly Ile Arg Glu Lys Leu Ala Gly Met Glu Gln Ser Arg Ser Glu
    290                 295                 300
Thr Ala Arg Ser Val Ala Arg Asn Ile
305                 310
```

<210> 57
<211> 1203
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1203)

<400> 57

```
atg acc gtc atg agt acg acc ata tcc aca gcc ccg agc ggc gcc gcg        48
Met Thr Val Met Ser Thr Thr Ile Ser Thr Ala Pro Ser Gly Ala Ala
 1               5                  10                  15

ctt gcg ccg tct cgc ata gat atg cgg gcg ccg gag ccc ggg agt gcc        96
Leu Ala Pro Ser Arg Ile Asp Met Arg Ala Pro Glu Pro Gly Ser Ala
            20                  25                  30

ggc gaa ggc gcc ggt atc ctg gcg ccg gtg acg acg ctg gct ctg gcg       144
Gly Glu Gly Ala Gly Ile Leu Ala Pro Val Thr Thr Leu Ala Leu Ala
        35                  40                  45

gcg ggc cgg ccg gct ttg cca gcg tca ccg tcg ctg cgc acc gcg ccc       192
Ala Gly Arg Pro Ala Leu Pro Ala Ser Pro Ser Leu Arg Thr Ala Pro
    50                  55                  60

gtc ctg gat ccg cca gtg cgc gat ctc agc ccc gcc gac ttg gcc gac       240
Val Leu Asp Pro Pro Val Arg Asp Leu Ser Pro Ala Asp Leu Ala Asp
65                  70                  75                  80

ctg ctg cgc gtc ttg cga tcc agg gcg gtg gac ggg cag ttg gcc acg       288
Leu Leu Arg Val Leu Arg Ser Arg Ala Val Asp Gly Gln Leu Ala Thr
                85                  90                  95

gcg cgc gag aac ctg cag gat gcg caa gtc aag gcg aag cag aac acc       336
Ala Arg Glu Asn Leu Gln Asp Ala Gln Val Lys Ala Lys Gln Asn Thr
            100                 105                 110

cag gcc cag ctc gac aag ctg gac gca tgg ttt cgg aag gct gag gac       384
Gln Ala Gln Leu Asp Lys Leu Asp Ala Trp Phe Arg Lys Ala Glu Asp
        115                 120                 125

gcc gag agc aag ggc tgg ctg agc aag gtg ttc ggc tgg atc ggg aag       432
Ala Glu Ser Lys Gly Trp Leu Ser Lys Val Phe Gly Trp Ile Gly Lys
    130                 135                 140

gtg ctg gcg gtc gtg gca tcg gcc ctg gct gtg ggc ttt gct gcc gtc       480
Val Leu Ala Val Val Ala Ser Ala Leu Ala Val Gly Phe Ala Ala Val
145                 150                 155                 160

gcc agc gtg gtc acc ggc gcg gcg gcc acg ccc atg ctg gtg ctc agc       528
```

```
Ala Ser Val Val Thr Gly Ala Ala Ala Thr Pro Met Leu Val Leu Ser
                  165                 170                 175

ggc atg gca ttg gtc agc gcc gtg aca tcg ctg gcc gac cag ata tcg      576
Gly Met ala Leu Val Ser Ala Val Thr Ser Leu Ala Asp Gln Ile Ser
            180                 185                 190

cga gag gcg gga ggg ccg cct atc agc ctg ggc ggg ttt ctc tcc ggg      624
Arg Glu Ala Gly Gly Pro Pro Ile Ser Leu Gly Gly Phe Leu Ser Gly
        195                 200                 205

ctg gcc gga cgt ctg ctg aca gcg ttg ggg gtg gat cag tcg cag gcc      672
Leu Ala Gly Arg Leu Leu Thr Ala Leu Gly Val Asp Gln Ser Gln Ala
    210                 215                 220

gac caa att gcc aag atc gtc gcc ggc ctg gcc gtg ccc gcc gtc ttg      720
Asp Gln Ile Ala Lys Ile Val Ala Gly Leu Ala Val Pro Ala Val Leu
225                 230                 235                 240

ctg atc gaa ccc cag atg ctg ggc gaa atg gcc gaa ggc gtg gcc agg      768
Leu Ile Glu Pro Gln Met Leu Gly Glu Met ala Glu Gly Val Ala Arg
                245                 250                 255

ctg gcg ggc gcc ggc gat gcc acc gcg gga tac ata gcc atg gcg atg      816
Leu Ala Gly Ala Gly Asp Ala Thr Ala Gly Tyr Ile Ala Met ala Met
            260                 265                 270

tcc atc gtg gcg gcg atc gcg gtc gcc gcg atc aat gcc gcc ggt acg      864
Ser Ile Val Ala Ala Ile Ala Val Ala Ala Ile Asn Ala Ala Gly Thr
        275                 280                 285

gcc ggc gcg ggc agc gcc tcg gcg atc agg ggt gcc tgg gat cgg gcc      912
Ala Gly Ala Gly Ser Ala Ser Ala Ile Arg Gly Ala Trp Asp Arg Ala
    290                 295                 300

gcc gcg gta gcc acc cag gtc ctt cag ggg ggt acg gca gtg gcg caa      960
Ala Ala Val Ala Thr Gln Val Leu Gln Gly Gly Thr Ala Val Ala Gln
305                 310                 315                 320

ggc ggc gtc ggc gtg tcg atg gca gtc gat cgc aaa cag gcc gat ctc     1008
Gly Gly Val Gly Val Ser Met ala Val Asp Arg Lys Gln Ala Asp Leu
                325                 330                 335

ctg gtc gcc gac aag gcg gat ctg gcg gcg agc ctg aca aaa ctg cgg     1056
Leu Val Ala Asp Lys Ala Asp Leu Ala Ala Ser Leu Thr Lys Leu Arg
            340                 345                 350

gcg gcc atg gag cgt gag gcg gac gat atc aag aag atc ctg gct caa     1104
Ala Ala Met Glu Arg Glu Ala Asp Asp Ile Lys Lys Ile Leu Ala Gln
        355                 360                 365

ttc gac gcg gcc tat cac atg atc gcg cag atg atc agc gac atg gcg     1152
Phe Asp Ala Ala Tyr His Met Ile Ala Gln Met Ile Ser Asp Met ala
    370                 375                 380

agc acg cac agc cag gtc agc gcc aac ctc gga cgg cgc cag gcg gtg     1200
Ser Thr His Ser Gln Val Ser Ala Asn Leu Gly Arg Arg Gln Ala Val
385                 390                 395                 400

tag                                                                1203
```

<210> 58
<211> 400
<212> PRT
<213> Bordetella pertussis

<400> 58

```
Met Thr Val Met Ser Thr Thr Ile Ser Thr Ala Pro Ser Gly Ala Ala
 1               5                  10                  15
Leu Ala Pro Ser Arg Ile Asp Met Arg Ala Pro Glu Pro Gly Ser Ala
            20                  25                  30
Gly Glu Gly Ala Gly Ile Leu Ala Pro Val Thr Thr Leu Ala Leu Ala
        35                  40                  45
Ala Gly Arg Pro Ala Leu Pro Ala Ser Pro Ser Leu Arg Thr Ala Pro
    50                  55                  60
Val Leu Asp Pro Pro Val Arg Asp Leu Ser Pro Ala Asp Leu Ala Asp
65                  70                  75                  80
Leu Leu Arg Val Leu Arg Ser Arg Ala Val Asp Gly Gln Leu Ala Thr
                85                  90                  95
Ala Arg Glu Asn Leu Gln Asp Ala Gln Val Lys Ala Lys Gln Asn Thr
            100                 105                 110
Gln Ala Gln Leu Asp Lys Leu Asp Ala Trp Phe Arg Lys Ala Glu Asp
        115                 120                 125
Ala Glu Ser Lys Gly Trp Leu Ser Lys Val Phe Gly Trp Ile Gly Lys
    130                 135                 140
Val Leu Ala Val Val Ala Ser Ala Leu Ala Val Gly Phe Ala Ala Val
145                 150                 155                 160
Ala Ser Val Val Thr Gly Ala Ala Ala Thr Pro Met Leu Val Leu Ser
                165                 170                 175
Gly Met ala Leu Val Ser Ala Val Thr Ser Leu Ala Asp Gln Ile Ser
            180                 185                 190
Arg Glu Ala Gly Gly Pro Pro Ile Ser Leu Gly Gly Phe Leu Ser Gly
        195                 200                 205
Leu Ala Gly Arg Leu Leu Thr Ala Leu Gly Val Asp Gln Ser Gln Ala
    210                 215                 220
Asp Gln Ile Ala Lys Ile Val Ala Gly Leu Ala Val Pro Ala Val Leu
225                 230                 235                 240
Leu Ile Glu Pro Gln Met Leu Gly Glu Met ala Glu Gly Val Ala Arg
                245                 250                 255
Leu Ala Gly Ala Gly Asp Ala Thr Ala Gly Tyr Ile Ala Met ala Met
            260                 265                 270
Ser Ile Val Ala Ala Ile Ala Val Ala Ala Ile Asn Ala Ala Gly Thr
        275                 280                 285
Ala Gly Ala Gly Ser Ala Ser Ala Ile Arg Gly Ala Trp Asp Arg Ala
    290                 295                 300
Ala Ala Val Ala Thr Gln Val Leu Gln Gly Gly Thr Ala Val Ala Gln
305                 310                 315                 320
Gly Gly Val Gly Val Ser Met ala Val Asp Arg Lys Gln Ala Asp Leu
                325                 330                 335
Leu Val Ala Asp Lys Ala Asp Leu Ala Ala Ser Leu Thr Lys Leu Arg
            340                 345                 350
Ala Ala Met Glu Arg Glu Ala Asp Asp Ile Lys Lys Ile Leu Ala Gln
        355                 360                 365
Phe Asp Ala Ala Tyr His Met Ile Ala Gln Met Ile Ser Asp Met ala
    370                 375                 380
Ser Thr His Ser Gln Val Ser Ala Asn Leu Gly Arg Arg Gln Ala Val
385                 390                 395                 400
```

<210> 59
<211> 462

<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(462)

<400> 59

```
atg act gtt cat gac gac gcg gcc gcg gcg ctg cgc gcc cgg ctg gat        48
Met Thr Val His Asp Asp Ala Ala Ala Ala Leu Arg Ala Arg Leu Asp
 1               5                  10                  15

gcg ttg ccg ggc agc cgg cgc ctg aca gcc gag caa ttg gaa gtg att        96
Ala Leu Pro Gly Ser Arg Arg Leu Thr Ala Glu Gln Leu Glu Val Ile
            20                  25                  30

tac gcg atg gcg tat gcg cac gtc gcc agg tgc gag tac ggc aag gcg       144
Tyr Ala Met ala Tyr Ala His Val Ala Arg Cys Glu Tyr Gly Lys Ala
        35                  40                  45

ctg ccc att ttc gcc ttc ctc gcg cag tac ggc ccc acg cgc aag cac       192
Leu Pro Ile Phe Ala Phe Leu Ala Gln Tyr Gly Pro Thr Arg Lys His
    50                  55                  60

tac tgg gcc ggc ctg gcg cta tgc ctg cag aag acc gac cgt ccc gac       240
Tyr Trp Ala Gly Leu Ala Leu Cys Leu Gln Lys Thr Asp Arg Pro Asp
65                  70                  75                  80

gag gcg cgc aat atc tat gcg ttg atc ctc acg ctc tat cca gat tcc       288
Glu Ala Arg Asn Ile Tyr Ala Leu Ile Leu Thr Leu Tyr Pro Asp Ser
                85                  90                  95

gcg gat gcc gtg ttg cgc acg gcc gag tgc gag ctg gcg ttg ggt gag       336
Ala Asp Ala Val Leu Arg Thr Ala Glu Cys Glu Leu Ala Leu Gly Glu
            100                 105                 110

aac gaa cgg gcg cag gcg gcc ctg ttc ggc gca atc gcc atc gat gca       384
Asn Glu Arg Ala Gln Ala Ala Leu Phe Gly Ala Ile Ala Ile Asp Ala
        115                 120                 125

gaa agt ggg cag cca ggt ccg gtc tcg cac cgt gcg cgc gct ttg ctc       432
Glu Ser Gly Gln Pro Gly Pro Val Ser His Arg Ala Arg Ala Leu Leu
    130                 135                 140

gat ctt att tca gtt tca cat ccg gag taa                              462
Asp Leu Ile Ser Val Ser His Pro Glu  *
145                 150
```

<210> 60
<211> 153
<212> PRT
<213> Bordetella pertussis

<400> 60

```
Met Thr Val His Asp Asp Ala Ala Ala Ala Leu Arg Ala Arg Leu Asp
 1               5                  10                  15
Ala Leu Pro Gly Ser Arg Arg Leu Thr Ala Glu Gln Leu Glu Val Ile
            20                  25                  30
Tyr Ala Met ala Tyr Ala His Val Ala Arg Cys Glu Tyr Gly Lys Ala
```

```
          35                  40                  45
Leu Pro Ile Phe Ala Phe Leu Ala Gln Tyr Gly Pro Thr Arg Lys His
      50                  55                  60
Tyr Trp Ala Gly Leu Ala Leu Cys Leu Gln Lys Thr Asp Arg Pro Asp
65                  70                  75                  80
Glu Ala Arg Asn Ile Tyr Ala Leu Ile Leu Thr Leu Tyr Pro Asp Ser
                  85                  90                  95
Ala Asp Ala Val Leu Arg Thr Ala Glu Cys Glu Leu Ala Leu Gly Glu
             100                 105                 110
Asn Glu Arg Ala Gln Ala Ala Leu Phe Gly Ala Ile Ala Ile Asp Ala
         115                 120                 125
Glu Ser Gly Gln Pro Gly Pro Val Ser His Arg Ala Arg Ala Leu Leu
     130                 135                 140
Asp Leu Ile Ser Val Ser His Pro Glu
145                 150
```

<210> 61
<211> 522
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(522)

<400> 61

```
atg cac tca gac tca ggt tca gat tca ggc tca gac tca ggc tca ggc      48
Met His Ser Asp Ser Gly Ser Asp Ser Gly Ser Asp Ser Gly Ser Gly
 1               5                  10                  15

tca ccc atg gtc tcg tcg ata cat cca tcg gaa ccg ata cag ccg atg      96
Ser Pro Met Val Ser Ser Ile His Pro Ser Glu Pro Ile Gln Pro Met
            20                  25                  30

gag cat gtg ctc gag gag gcc gac gcc cgc ctg ctt acc gaa gtg ggt     144
Glu His Val Leu Glu Glu Ala Asp Ala Arg Leu Leu Thr Glu Val Gly
        35                  40                  45

ttt ctg gcg gcg gcc gtc agc gat ctg acg cgc gcg gac gcc att ttc     192
Phe Leu Ala Ala Ala Val Ser Asp Leu Thr Arg Ala Asp Ala Ile Phe
    50                  55                  60

aat gca ttg caa cgt gta cgg ccg ggc cgg acg cat ccc tgc atc ggc     240
Asn Ala Leu Gln Arg Val Arg Pro Gly Arg Thr His Pro Cys Ile Gly
65                  70                  75                  80

ctg gcg gtc gcc cgc atg aac gcc ggg ctg ccc gac gaa gcc gcc gag     288
Leu Ala Val Ala Arg Met Asn Ala Gly Leu Pro Asp Glu Ala Ala Glu
                85                  90                  95

atc ctg gcg aat ttc cag ccg gca cag ccg gag gac cgc tcg gaa ctg     336
Ile Leu Ala Asn Phe Gln Pro Ala Gln Pro Glu Asp Arg Ser Glu Leu
            100                 105                 110

gac gcc tgg tgc ggg ttc gct ctg ttg ctg gct ggc cgc tcg gac gag     384
Asp Ala Trp Cys Gly Phe Ala Leu Leu Leu Ala Gly Arg Ser Asp Glu
        115                 120                 125

gcg cgc cgc atg ctg cag cga gcc atc gat gcg ggt ggc gag gcg gca     432
Ala Arg Arg Met Leu Gln Arg Ala Ile Asp Ala Gly Gly Glu Ala Ala
```

```
        130                 135                 140

agg ctg gcg cag gtc gtg ttg gac agc gga ccc gcc atg atg cgg ccc      480
Arg Leu Ala Gln Val Val Leu Asp Ser Gly Pro Ala Met Met Arg Pro
145                 150                 155                 160

gcg ccg ttg cag tcc gag cca tta cct gga gct cct gga tga            522
Ala Pro Leu Gln Ser Glu Pro Leu Pro Gly Ala Pro Gly  *
                165                 170
```

<210> 62
<211> 173
<212> PRT
<213> Bordetella pertussis

<400> 62

```
Met His Ser Asp Ser Gly Ser Asp Ser Gly Ser Asp Ser Gly Ser Gly
 1               5                  10                  15
Ser Pro Met Val Ser Ser Ile His Pro Ser Glu Pro Ile Gln Pro Met
            20                  25                  30
Glu His Val Leu Glu Glu Ala Asp Ala Arg Leu Leu Thr Glu Val Gly
        35                  40                  45
Phe Leu Ala Ala Ala Val Ser Asp Leu Thr Arg Ala Asp Ala Ile Phe
    50                  55                  60
Asn Ala Leu Gln Arg Val Arg Pro Gly Arg Thr His Pro Cys Ile Gly
65                  70                  75                  80
Leu Ala Val Ala Arg Met Asn Ala Gly Leu Pro Asp Glu Ala Ala Glu
                85                  90                  95
Ile Leu Ala Asn Phe Gln Pro Ala Gln Pro Glu Asp Arg Ser Glu Leu
            100                 105                 110
Asp Ala Trp Cys Gly Phe Ala Leu Leu Leu Ala Gly Arg Ser Asp Glu
        115                 120                 125
Ala Arg Arg Met Leu Gln Arg Ala Ile Asp Ala Gly Gly Glu Ala Ala
    130                 135                 140
Arg Leu Ala Gln Val Val Leu Asp Ser Gly Pro Ala Met Met Arg Pro
145                 150                 155                 160
Ala Pro Leu Gln Ser Glu Pro Leu Pro Gly Ala Pro Gly
                165                 170
```

<210> 63
<211> 180
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(180)

<400> 63

```
atg gtt gcg cga ggt cct tgc gca cca cct ggt gta gtc ccc ata cgt       48
Met Val Ala Arg Gly Pro Cys Ala Pro Pro Gly Val Val Pro Ile Arg
 1               5                  10                  15

cgc gtc gta agt ttt acg gaa gac gtt cag cgc gtt gta tct gaa tgc       96
Arg Val Val Ser Phe Thr Glu Asp Val Gln Arg Val Val Ser Glu Cys
            20                  25                  30

tcc ggt agc gac cgc gat ccg aca tta gtt tcg gaa gtt aac aac tgc      144
Ser Gly Ser Asp Arg Asp Pro Thr Leu Val Ser Glu Val Asn Asn Cys
```

```
          35                  40                  45

cgg ata aaa ccg gca gtc agt atc ttg gtg acg tga                    180
Arg Ile Lys Pro Ala Val Ser Ile Leu Val Thr  *
 50                  55

      <210> 64
      <211> 59
      <212> PRT
      <213> Bordetella pertussis

      <400> 64
Met Val Ala Arg Gly Pro Cys Ala Pro Pro Gly Val Val Pro Ile Arg
 1               5                  10                  15
Arg Val Val Ser Phe Thr Glu Asp Val Gln Arg Val Val Ser Glu Cys
            20                  25                  30
Ser Gly Ser Asp Arg Asp Pro Thr Leu Val Ser Glu Val Asn Asn Cys
        35                  40                  45
Arg Ile Lys Pro Ala Val Ser Ile Leu Val Thr
    50                  55

      <210> 65
      <211> 975
      <212> DNA
      <213> Bordetella pertussis

      <220>
      <221> CDS
      <222> (1)...(975)

      <400> 65
atg cgg ttt cga gca ggt tat agc cgt tat caa gcc cgc tca ggc cat     48
Met Arg Phe Arg Ala Gly Tyr Ser Arg Tyr Gln Ala Arg Ser Gly His
 1               5                  10                  15

ggg gac cgc cca ccc ccc gca cag gcg cgc gtc cag acg gta ctc ctg     96
Gly Asp Arg Pro Pro Pro Ala Gln Ala Arg Val Gln Thr Val Leu Leu
            20                  25                  30

cac gga ctc tcc gcg ttg acg gcg caa gtt gcg cag cgc ttc gaa atg    144
His Gly Leu Ser Ala Leu Thr Ala Gln Val Ala Gln Arg Phe Glu Met
        35                  40                  45

gcg cgc cac cgg atg gct ggc ccc ggt cgc acg aca ggc cac cac cat    192
Ala Arg His Arg Met ala Gly Pro Gly Arg Thr Thr Gly His His His
    50                  55                  60

ttc cag ctc gag gcc cag cgt atg gcc gac act ttg cgc agc gtt caa    240
Phe Gln Leu Glu Ala Gln Arg Met ala Asp Thr Leu Arg Ser Val Gln
 65                  70                  75                  80

ggc gag cct cgg tgg ccg gac ggg agc gag gcc tgc atg ccg tcg ggt    288
Gly Glu Pro Arg Trp Pro Asp Gly Ser Glu Ala Cys Met Pro Ser Gly
                85                  90                  95

ttg tca tgc cgg cat gga acc gaa gag ccg aaa gcg tca cac agt gca    336
Leu Ser Cys Arg His Gly Thr Glu Glu Pro Lys Ala Ser His Ser Ala
            100                 105                 110
```

```
tat tcc atg ttc ccg ctc cgt aga acg cga tac acc caa gga ttc gag      384
Tyr Ser Met Phe Pro Leu Arg Arg Thr Arg Tyr Thr Gln Gly Phe Glu
        115                 120                 125

aca acg gcc cac cgc atg aac ttc cag atc cca ccc gct tta cct gct      432
Thr Thr Ala His Arg Met Asn Phe Gln Ile Pro Pro Ala Leu Pro Ala
    130                 135                 140

ttg gag ctt gat gtc ttt gcg cgc gcc gcc agc caa gga gag acc cta      480
Leu Glu Leu Asp Val Phe Ala Arg Ala Ala Ser Gln Gly Glu Thr Leu
145                 150                 155                 160

tat gtc acc aaa gca ggc gag cag ttc cag gtc atc gca tcc ggc acg      528
Tyr Val Thr Lys Ala Gly Glu Gln Phe Gln Val Ile Ala Ser Gly Thr
                165                 170                 175

acg ccg tca ggg cgc aac gta tcc tgg gtc gcc acc gac gag gac acg      576
Thr Pro Ser Gly Arg Asn Val Ser Trp Val Ala Thr Asp Glu Asp Thr
            180                 185                 190

ctt gtc atg ttt tcc agc gcg ctg gcg ctg gcc tac ggc acg gga atc      624
Leu Val Met Phe Ser Ser Ala Leu Ala Leu Ala Tyr Gly Thr Gly Ile
        195                 200                 205

gcc cgc gcc gtc gcc aag gag ctc gat ctg cac gcg gtc ccg acg aca      672
Ala Arg Ala Val Ala Lys Glu Leu Asp Leu His Ala Val Pro Thr Thr
    210                 215                 220

tcg ctg tcg gcg cgc gtc gtc acg cga gcg gtc gac atg gcg gaa acc      720
Ser Leu Ser Ala Arg Val Val Thr Arg Ala Val Asp Met ala Glu Thr
225                 230                 235                 240

tca cgc cac gcc ctg cag ggc gtg gat ttc ctt acc ttc ctg tcc tgg      768
Ser Arg His Ala Leu Gln Gly Val Asp Phe Leu Thr Phe Leu Ser Trp
                245                 250                 255

tcg gcc cgc gcc gac gcc gcc ggc ttc cga cag gtc tgt cac gac acc      816
Ser Ala Arg Ala Asp Ala Ala Gly Phe Arg Gln Val Cys His Asp Thr
            260                 265                 270

ggt gtc tct ccc gat cag ata tcc gga acg ttg cgc gcc acg atc gac      864
Gly Val Ser Pro Asp Gln Ile Ser Gly Thr Leu Arg Ala Thr Ile Asp
        275                 280                 285

gaa agc atg cag cag cgc ttc gca tcc gcc gca caa tca ggt aag gcg      912
Glu Ser Met Gln Gln Arg Phe Ala Ser Ala Ala Gln Ser Gly Lys Ala
    290                 295                 300

ccg gta tcc gcc cat acg gcg caa gaa tgg ttg cgc gag gtc ctt gcg      960
Pro Val Ser Ala His Thr Ala Gln Glu Trp Leu Arg Glu Val Leu Ala
305                 310                 315                 320

cac cac ctg gtg tag                                                  975
His His Leu Val  *
```

<210> 66
<211> 324
<212> PRT

<213> Bordetella pertussis

<400> 66

Met Arg Phe Arg Ala Gly Tyr Ser Arg Tyr Gln Ala Arg Ser Gly His
1 5 10 15
Gly Asp Arg Pro Pro Pro Ala Gln Ala Arg Val Gln Thr Val Leu Leu
20 25 30
His Gly Leu Ser Ala Leu Thr Ala Gln Val Ala Gln Arg Phe Glu Met
35 40 45
Ala Arg His Arg Met ala Gly Pro Gly Arg Thr Thr Gly His His His
50 55 60
Phe Gln Leu Glu Ala Gln Arg Met ala Asp Thr Leu Arg Ser Val Gln
65 70 75 80
Gly Glu Pro Arg Trp Pro Asp Gly Ser Glu Ala Cys Met Pro Ser Gly
85 90 95
Leu Ser Cys Arg His Gly Thr Glu Glu Pro Lys Ala Ser His Ser Ala
100 105 110
Tyr Ser Met Phe Pro Leu Arg Arg Thr Arg Tyr Thr Gln Gly Phe Glu
115 120 125
Thr Thr Ala His Arg Met Asn Phe Gln Ile Pro Pro Ala Leu Pro Ala
130 135 140
Leu Glu Leu Asp Val Phe Ala Arg Ala Ala Ser Gln Gly Glu Thr Leu
145 150 155 160
Tyr Val Thr Lys Ala Gly Glu Gln Phe Gln Val Ile Ala Ser Gly Thr
165 170 175
Thr Pro Ser Gly Arg Asn Val Ser Trp Val Ala Thr Asp Glu Asp Thr
180 185 190
Leu Val Met Phe Ser Ser Ala Leu Ala Leu Ala Tyr Gly Thr Gly Ile
195 200 205
Ala Arg Ala Val Ala Lys Glu Leu Asp Leu His Ala Val Pro Thr Thr
210 215 220
Ser Leu Ser Ala Arg Val Val Thr Arg Ala Val Asp Met ala Glu Thr
225 230 235 240
Ser Arg His Ala Leu Gln Gly Val Asp Phe Leu Thr Phe Leu Ser Trp
245 250 255
Ser Ala Arg Ala Asp Ala Ala Gly Phe Arg Gln Val Cys His Asp Thr
260 265 270
Gly Val Ser Pro Asp Gln Ile Ser Gly Thr Leu Arg Ala Thr Ile Asp
275 280 285
Glu Ser Met Gln Gln Arg Phe Ala Ser Ala Ala Gln Ser Gly Lys Ala
290 295 300
Pro Val Ser Ala His Thr Ala Gln Glu Trp Leu Arg Glu Val Leu Ala
305 310 315 320
His His Leu Val

<210> 67
<211> 1146
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1146)

<400> 67

atg ctg atc aac gcg gcc gag cac ccc gcc gcc agc ctg gat gcc gac 48
Met Leu Ile Asn Ala Ala Glu His Pro Ala Ala Ser Leu Asp Ala Asp
1 5 10 15

```
tgg tac cgg cga gtg cgg gtg ccg cgg ccc atc tac gag gaa ctc gtc        96
Trp Tyr Arg Arg Val Arg Val Pro Arg Pro Ile Tyr Glu Glu Leu Val
                20                  25                  30

ggc cag cga ggc tgg ctg cac cgg atc ggg ata gac gcc aag gca cag       144
Gly Gln Arg Gly Trp Leu His Arg Ile Gly Ile Asp Ala Lys Ala Gln
            35                  40                  45

aac agc ccc tgc acg tcg gtt ccc gtg gcc atc gcc gcg cgc tgc ctg       192
Asn Ser Pro Cys Thr Ser Val Pro Val Ala Ile Ala Ala Arg Cys Leu
        50                  55                  60

aac gtc gtg ctg gcg ctg gct ccc gcg cag atc gcc atg ttc gcc aac       240
Asn Val Val Leu Ala Leu Ala Pro Ala Gln Ile Ala Met Phe Ala Asn
    65                  70                  75                  80

agc ccg ctg gag gca ggg cgg gtg acc ggt ctc aag gaa aac cgc ctg       288
Ser Pro Leu Glu Ala Gly Arg Val Thr Gly Leu Lys Glu Asn Arg Leu
                85                  90                  95

acc ctg tgg ccg cgc atg ttc cga ggc gcg cgc tac ctg ggc gac gac       336
Thr Leu Trp Pro Arg Met Phe Arg Gly Ala Arg Tyr Leu Gly Asp Asp
            100                 105                 110

ctg ctg cat cgc ctg cct gca agg ccg ttt cgc gat ctc ggc gat tat       384
Leu Leu His Arg Leu Pro Ala Arg Pro Phe Arg Asp Leu Gly Asp Tyr
        115                 120                 125

ttc cgc tgg atg ttc ggc gga ttg acc gcc agc cgg gcg cta ccg ccg       432
Phe Arg Trp Met Phe Gly Gly Leu Thr Ala Ser Arg Ala Leu Pro Pro
    130                 135                 140

ggc gac gct tgc gac tac aag aac gcc gat gtg gcc tgc ctg gtg gga       480
Gly Asp Ala Cys Asp Tyr Lys Asn Ala Asp Val Ala Cys Leu Val Gly
145                 150                 155                 160

gcc cct tcg ctg gca gag ttc ctg tat gcg ggc gcg tgg tcc gcg cga       528
Ala Pro Ser Leu Ala Glu Phe Leu Tyr Ala Gly Ala Trp Ser Ala Arg
                165                 170                 175

aac ctg aat gat ggc ggt tcc gtg cgt ctg gcc gcg cgc agc gaa cat       576
Asn Leu Asn Asp Gly Gly Ser Val Arg Leu Ala Ala Arg Ser Glu His
            180                 185                 190

ttc gtc tat tcg cag ttc gcg cag ttc ctg gac gcg cgt tgg cgc tac       624
Phe Val Tyr Ser Gln Phe Ala Gln Phe Leu Asp Ala Arg Trp Arg Tyr
        195                 200                 205

agg atg ccg att gtc ccc gcc ttg ccg gcg ctg ttg cga gcc tgg gac       672
Arg Met Pro Ile Val Pro Ala Leu Pro Ala Leu Leu Arg Ala Trp Asp
    210                 215                 220

agg cag ggc ggc ctg gaa gcg ctg ttc gag cag gcc ggc gcg caa ggc       720
Arg Gln Gly Gly Leu Glu Ala Leu Phe Glu Gln Ala Gly Ala Gln Gly
225                 230                 235                 240

tac atc gag ggg cgc gcg ccg ggc gcg gta ttt gcc gat gcc gac ttg       768
Tyr Ile Glu Gly Arg Ala Pro Gly Ala Val Phe Ala Asp Ala Asp Leu
                245                 250                 255
```

```
ctg agc tca gcc ggc gat gca gtc gcg gcc agt gcg ccg atg gcg gcg      816
Leu Ser Ser Ala Gly Asp Ala Val Ala Ala Ser Ala Pro Met ala Ala
            260                 265                 270

tcg gcg ctg caa ttg ggg ctg ttg cgc aat ctg cac gac gcc gag gcc      864
Ser Ala Leu Gln Leu Gly Leu Leu Arg Asn Leu His Asp Ala Glu Ala
        275                 280                 285

ctg gtg agg cga tgg ggc tgg ctg cgc ttg cgt gcg ttg cgc gat cgg      912
Leu Val Arg Arg Trp Gly Trp Leu Arg Leu Arg Ala Leu Arg Asp Arg
    290                 295                 300

gcc atc gct ttg gcg ttg gac gat gcg cag gtg cgc tgc ctt tgc caa      960
Ala Ile Ala Leu Ala Leu Asp Asp Ala Gln Val Arg Cys Leu Cys Gln
305                 310                 315                 320

cag gtc gtg gcg gta gcc gaa ggc ggg ctg gcc ggc gac gag cag caa     1008
Gln Val Val Ala Val Ala Glu Gly Gly Leu Ala Gly Asp Glu Gln Gln
                325                 330                 335

tgg ctc gat tat gtg cgt tac gtg gtg gaa acc ggc gag acc gcc gcg     1056
Trp Leu Asp Tyr Val Arg Tyr Val Val Glu Thr Gly Glu Thr Ala Ala
            340                 345                 350

gac cgc atg ctg cgc ttg tgg cgc cag gcg cgc ggc acg cct gag atg     1104
Asp Arg Met Leu Arg Leu Trp Arg Gln Ala Arg Gly Thr Pro Glu Met
        355                 360                 365

cgc cgc gca cag gcg tgc cgg cag cgc gcg gtg ctg tcc tag             1146
Arg Arg Ala Gln Ala Cys Arg Gln Arg Ala Val Leu Ser  *
    370                 375                 380
```

<210> 68
<211> 381
<212> PRT
<213> Bordetella pertussis

<400> 68

```
Met Leu Ile Asn Ala Ala Glu His Pro Ala Ala Ser Leu Asp Ala Asp
 1               5                  10                  15
Trp Tyr Arg Arg Val Arg Val Pro Arg Pro Ile Tyr Glu Glu Leu Val
            20                  25                  30
Gly Gln Arg Gly Trp Leu His Arg Ile Gly Ile Asp Ala Lys Ala Gln
        35                  40                  45
Asn Ser Pro Cys Thr Ser Val Pro Val Ala Ile Ala Ala Arg Cys Leu
    50                  55                  60
Asn Val Val Leu Ala Leu Ala Pro Ala Gln Ile Ala Met Phe Ala Asn
65                  70                  75                  80
Ser Pro Leu Glu Ala Gly Arg Val Thr Gly Leu Lys Glu Asn Arg Leu
                85                  90                  95
Thr Leu Trp Pro Arg Met Phe Arg Gly Ala Arg Tyr Leu Gly Asp Asp
            100                 105                 110
Leu Leu His Arg Leu Pro Ala Arg Pro Phe Arg Asp Leu Gly Asp Tyr
        115                 120                 125
Phe Arg Trp Met Phe Gly Gly Leu Thr Ala Ser Arg Ala Leu Pro Pro
    130                 135                 140
Gly Asp Ala Cys Asp Tyr Lys Asn Ala Asp Val Ala Cys Leu Val Gly
145                 150                 155                 160
Ala Pro Ser Leu Ala Glu Phe Leu Tyr Ala Gly Ala Trp Ser Ala Arg
```

```
                    165                 170                 175
Asn Leu Asn Asp Gly Gly Ser Val Arg Leu Ala Ala Arg Ser Glu His
                180                 185                 190
Phe Val Tyr Ser Gln Phe Ala Gln Phe Leu Asp Ala Arg Trp Arg Tyr
            195                 200                 205
Arg Met Pro Ile Val Pro Ala Leu Pro Ala Leu Leu Arg Ala Trp Asp
        210                 215                 220
Arg Gln Gly Gly Leu Glu Ala Leu Phe Glu Gln Ala Gly Ala Gln Gly
225                 230                 235                 240
Tyr Ile Glu Gly Arg Ala Pro Gly Ala Val Phe Ala Asp Ala Asp Leu
                245                 250                 255
Leu Ser Ser Ala Gly Asp Ala Val Ala Ala Ser Ala Pro Met ala Ala
            260                 265                 270
Ser Ala Leu Gln Leu Gly Leu Leu Arg Asn Leu His Asp Ala Glu Ala
        275                 280                 285
Leu Val Arg Arg Trp Gly Trp Leu Arg Leu Arg Ala Leu Arg Asp Arg
    290                 295                 300
Ala Ile Ala Leu Ala Leu Asp Asp Ala Gln Val Arg Cys Leu Cys Gln
305                 310                 315                 320
Gln Val Val Ala Val Ala Glu Gly Gly Leu Ala Gly Asp Glu Gln Gln
                325                 330                 335
Trp Leu Asp Tyr Val Arg Tyr Val Val Glu Thr Gly Glu Thr Ala Ala
            340                 345                 350
Asp Arg Met Leu Arg Leu Trp Arg Gln Ala Arg Gly Thr Pro Glu Met
        355                 360                 365
Arg Arg Ala Gln Ala Cys Arg Gln Arg Ala Val Leu Ser
    370                 375                 380
```

<210> 69
<211> 1233
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1233)

<400> 69

```
atg acc ctt cgc ccc gga ata ctc gcc ccg ctc gcc ctg ctt ctt ggc      48
Met Thr Leu Arg Pro Gly Ile Leu Ala Pro Leu Ala Leu Leu Leu Gly
 1               5                  10                  15

ctc gcg tgg ccc atc gcg gcg gca tcc agc cct cct cct gtc gcg gcc      96
Leu Ala Trp Pro Ile Ala Ala Ala Ser Ser Pro Pro Pro Val Ala Ala
            20                  25                  30

ctg agc tcc ggc gtc gcc ctt acc tcg ccc cgc ctt ccg cct ccc tcc     144
Leu Ser Ser Gly Val Ala Leu Thr Ser Pro Arg Leu Pro Pro Pro Ser
        35                  40                  45

cat aca tcc ggc cgg aaa tgg cgc atc ggt tat gtg ggt agc ggc gag     192
His Thr Ser Gly Arg Lys Trp Arg Ile Gly Tyr Val Gly Ser Gly Glu
    50                  55                  60

tac gag gag tat ccg cgc acg ctc tac gcg atc gcg cgc gca ttg caa     240
Tyr Glu Glu Tyr Pro Arg Thr Leu Tyr Ala Ile Ala Arg Ala Leu Gln
65                  70                  75                  80

caa ctc gga tgg ctg cgt atc gac gac atg ccc gag ata acc gat atg     288
Gln Leu Gly Trp Leu Arg Ile Asp Asp Met Pro Glu Ile Thr Asp Met
```

```
                    85                  90                  95

cga aag gcc tgg ctt tac ctg gcc acg cat gcc cgc agc aac tac atc      336
Arg Lys Ala Trp Leu Tyr Leu Ala Thr His Ala Arg Ser Asn Tyr Ile
            100                 105                 110

gag ttc gtg ccc gat gcg tgg tgg cag ccc ggc aac ttc gac acc gcc      384
Glu Phe Val Pro Asp Ala Trp Trp Gln Pro Gly Asn Phe Asp Thr Ala
        115                 120                 125

ttg cgg cct gcc gtg cgc gaa gcc gtt gcg gca cgc ctg cat ggc gcc      432
Leu Arg Pro Ala Val Arg Glu Ala Val Ala Ala Arg Leu His Gly Ala
    130                 135                 140

aag gac atc gac ctg atc atc gcc atg ggt acc tgg gct gga cag gac      480
Lys Asp Ile Asp Leu Ile Ile Ala Met Gly Thr Trp Ala Gly Gln Asp
145                 150                 155                 160

atg gtc gaa ctg ggc acg ccg gta ccc acc gtg gtc gtc tcg tcg acc      528
Met Val Glu Leu Gly Thr Pro Val Pro Thr Val Val Val Ser Ser Thr
                165                 170                 175

gac ccg ata agc gcc cgg atc ata ccc agt gcg gcc gac agc ggc cag      576
Asp Pro Ile Ser Ala Arg Ile Ile Pro Ser Ala Ala Asp Ser Gly Gln
            180                 185                 190

gac aac ctg cat gcc cgg gta cag ccc gac cac tac cag cgg cag atc      624
Asp Asn Leu His Ala Arg Val Gln Pro Asp His Tyr Gln Arg Gln Ile
        195                 200                 205

cag ctg ctc cat gac atc gtg ccg ttc aag acg ctt gga ctg gtc tac      672
Gln Leu Leu His Asp Ile Val Pro Phe Lys Thr Leu Gly Leu Val Tyr
    210                 215                 220

gaa gac acc gaa gca ggt cgc acc tac gca gcc atc gat aag gtc gcc      720
Glu Asp Thr Glu Ala Gly Arg Thr Tyr Ala Ala Ile Asp Lys Val Ala
225                 230                 235                 240

gca cta atg ccg gca ttg gat ttc tcc gtc aag cgt tgc gac gca cgc      768
Ala Leu Met Pro Ala Leu Asp Phe Ser Val Lys Arg Cys Asp Ala Arg
                245                 250                 255

gcg acc ggc atc ccc atc gcc acg gca acc cag aac gtt ctg gct tgc      816
Ala Thr Gly Ile Pro Ile Ala Thr Ala Thr Gln Asn Val Leu Ala Cys
            260                 265                 270

tac cag aag ctg tcg agc gaa gtc gac gcc ttt tac gtc acc gag cac      864
Tyr Gln Lys Leu Ser Ser Glu Val Asp Ala Phe Tyr Val Thr Glu His
        275                 280                 285

cgg ggc atc acc tcg acg tcc gtc aag cag ctc gcc gcg ctg ctg cgc      912
Arg Gly Ile Thr Ser Thr Ser Val Lys Gln Leu Ala Ala Leu Leu Arg
    290                 295                 300

gcc gcc cgc gtg ccg agt ttc tcg atg caa ggc tcc gac gag gtc aag      960
Ala Ala Arg Val Pro Ser Phe Ser Met Gln Gly Ser Asp Glu Val Lys
305                 310                 315                 320

gcc ggc ctg ttg atg agc ctg gcc aag gcg gac tac tcc agc gta ggc     1008
Ala Gly Leu Leu Met Ser Leu Ala Lys Ala Asp Tyr Ser Ser Val Gly
```

```
                    325                 330                 335

atg ttc cac gcc cag acc att gcc cgc att ttc aat ggg gaa aag ccg      1056
Met Phe His Ala Gln Thr Ile Ala Arg Ile Phe Asn Gly Glu Lys Pro
            340                 345                 350

cgc agc atc agc cag gtc tgg aat gcc ccc gcc aag ata gcc atc aat      1104
Arg Ser Ile Ser Gln Val Trp Asn Ala Pro Ala Lys Ile Ala Ile Asn
        355                 360                 365

ctg gaa acg gcg cgg cgc atc ggc ttc gac cca ccg gtg gat att ctg      1152
Leu Glu Thr Ala Arg Arg Ile Gly Phe Asp Pro Pro Val Asp Ile Leu
    370                 375                 380

ctg gcg gcc gac gag gtg tac gaa gcg gag cac tga cag gcc tgg cca      1200
Leu Ala Ala Asp Glu Val Tyr Glu Ala Glu His  *  Gln Ala Trp Pro
385                 390                 395

acg aga cct ggc aag gaa tgt gcc gga tcc tag                         1233
Thr Arg Pro Gly Lys Glu Cys Ala Gly Ser  *
400                 405


      <210> 70
      <211> 409
      <212> PRT
      <213> Bordetella pertussis

      <400> 70
Met Thr Leu Arg Pro Gly Ile Leu Ala Pro Leu Ala Leu Leu Leu Gly
 1               5                  10                  15
Leu Ala Trp Pro Ile Ala Ala Ala Ser Ser Pro Pro Pro Val Ala Ala
            20                  25                  30
Leu Ser Ser Gly Val Ala Leu Thr Ser Pro Arg Leu Pro Pro Pro Ser
        35                  40                  45
His Thr Ser Gly Arg Lys Trp Arg Ile Gly Tyr Val Gly Ser Gly Glu
    50                  55                  60
Tyr Glu Glu Tyr Pro Arg Thr Leu Tyr Ala Ile Ala Arg Ala Leu Gln
65                  70                  75                  80
Gln Leu Gly Trp Leu Arg Ile Asp Asp Met Pro Glu Ile Thr Asp Met
                85                  90                  95
Arg Lys Ala Trp Leu Tyr Leu Ala Thr His Ala Arg Ser Asn Tyr Ile
            100                 105                 110
Glu Phe Val Pro Asp Ala Trp Trp Gln Pro Gly Asn Phe Asp Thr Ala
        115                 120                 125
Leu Arg Pro Ala Val Arg Glu Ala Val Ala Ala Arg Leu His Gly Ala
    130                 135                 140
Lys Asp Ile Asp Leu Ile Ile Ala Met Gly Thr Trp Ala Gly Gln Asp
145                 150                 155                 160
Met Val Glu Leu Gly Thr Pro Val Pro Thr Val Val Val Ser Ser Thr
                165                 170                 175
Asp Pro Ile Ser Ala Arg Ile Ile Pro Ser Ala Ala Asp Ser Gly Gln
            180                 185                 190
Asp Asn Leu His Ala Arg Val Gln Pro Asp His Tyr Gln Arg Gln Ile
        195                 200                 205
Gln Leu Leu His Asp Ile Val Pro Phe Lys Thr Leu Gly Leu Val Tyr
    210                 215                 220
Glu Asp Thr Glu Ala Gly Arg Thr Tyr Ala Ala Ile Asp Lys Val Ala
225                 230                 235                 240
Ala Leu Met Pro Ala Leu Asp Phe Ser Val Lys Arg Cys Asp Ala Arg
```

```
                    245                 250                 255
Ala Thr Gly Ile Pro Ile Ala Thr Ala Thr Gln Asn Val Leu Ala Cys
                260                 265                 270
Tyr Gln Lys Leu Ser Ser Glu Val Asp Ala Phe Tyr Val Thr Glu His
            275                 280                 285
Arg Gly Ile Thr Ser Thr Ser Val Lys Gln Leu Ala Ala Leu Leu Arg
    290                 295                 300
Ala Ala Arg Val Pro Ser Phe Ser Met Gln Gly Ser Asp Glu Val Lys
305                 310                 315                 320
Ala Gly Leu Leu Met Ser Leu Ala Lys Ala Asp Tyr Ser Ser Val Gly
                325                 330                 335
Met Phe His Ala Gln Thr Ile Ala Arg Ile Phe Asn Gly Glu Lys Pro
            340                 345                 350
Arg Ser Ile Ser Gln Val Trp Asn Ala Pro Ala Lys Ile Ala Ile Asn
        355                 360                 365
Leu Glu Thr Ala Arg Arg Ile Gly Phe Asp Pro Pro Val Asp Ile Leu
    370                 375                 380
Leu Ala Ala Asp Glu Val Tyr Glu Ala Glu His Gln Ala Trp Pro Thr
385                 390                 395                 400
Arg Pro Gly Lys Glu Cys Ala Gly Ser
                405
```

<210> 71
<211> 645
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(645)

<400> 71

```
atg gag cag ctg gat ctg ccg ctg gta gtg gtc ggg ctg tac ccg ggc        48
Met Glu Gln Leu Asp Leu Pro Leu Val Val Val Gly Leu Tyr Pro Gly
 1               5                  10                  15

atg cag gtt gtc ctg gcc gct gtc ggc cgc act ggg tat gat ccg ggc        96
Met Gln Val Val Leu Ala Ala Val Gly Arg Thr Gly Tyr Asp Pro Gly
            20                  25                  30

gct tat cgg gtc ggt cga cga gac gac cac ggt ggg tac cgg cgt gcc       144
Ala Tyr Arg Val Gly Arg Arg Asp Asp His Gly Gly Tyr Arg Arg Ala
        35                  40                  45

cag ttc gac cat gtc ctg tcc agc cca ggt acc cat ggc gat gat cag       192
Gln Phe Asp His Val Leu Ser Ser Pro Gly Thr His Gly Asp Asp Gln
    50                  55                  60

gtc gat gtc ctt ggc gcc atg cag gcg tgc cgc aac ggc ttc gcg cac       240
Val Asp Val Leu Gly Ala Met Gln Ala Cys Arg Asn Gly Phe Ala His
65                  70                  75                  80

ggc agg ccg caa ggc ggt gtc gaa gtt gcc ggg ctg cca cca cgc atc       288
Gly Arg Pro Gln Gly Gly Val Glu Val Ala Gly Leu Pro Pro Arg Ile
                85                  90                  95

ggg cac gaa ctc gat gta gtt gct gcg ggc atg cgt ggc cag gta aag       336
Gly His Glu Leu Asp Val Val Ala Ala Gly Met Arg Gly Gln Val Lys
            100                 105                 110
```

```
cca ggc ctt tcg cat atc ggt tat ctc ggg cat gtc gtc gat acg cag      384
Pro Gly Leu Ser His Ile Gly Tyr Leu Gly His Val Val Asp Thr Gln
        115                 120                 125

cca tcc gag ttg ttg caa tgc gcg cgc gat cgc gta gag cgt gcg cgg      432
Pro Ser Glu Leu Leu Gln Cys Ala Arg Asp Arg Val Glu Arg Ala Arg
    130                 135                 140

ata ctc ctc gta ctc gcc gct acc cac ata acc gat gcg cca ttt ccg      480
Ile Leu Leu Val Leu Ala Ala Thr His Ile Thr Asp Ala Pro Phe Pro
145                 150                 155                 160

gcc gga tgt atg gga ggg agg cgg aag gcg ggg cga ggt aag ggc gac      528
Ala Gly Cys Met Gly Gly Arg Arg Lys Ala Gly Arg Gly Lys Gly Asp
                165                 170                 175

gcc gga gct cag ggc cgc gac agg agg agg gct gga tgc cgc cgc gat      576
Ala Gly Ala Gln Gly Arg Asp Arg Arg Arg Ala Gly Cys Arg Arg Asp
            180                 185                 190

ggg cca cgc gag gcc aag aag cag ggc gag cgg ggc gag tat tcc ggg      624
Gly Pro Arg Glu Ala Lys Lys Gln Gly Glu Arg Gly Glu Tyr Ser Gly
        195                 200                 205

gcg aag ggt cat ggg cga tga                                         645
Ala Lys Gly His Gly Arg  *
    210
```

<210> 72
<211> 214
<212> PRT
<213> Bordetella pertussis

<400> 72

```
Met Glu Gln Leu Asp Leu Pro Leu Val Val Val Gly Leu Tyr Pro Gly
 1               5                  10                  15
Met Gln Val Val Leu Ala Ala Val Gly Arg Thr Gly Tyr Asp Pro Gly
            20                  25                  30
Ala Tyr Arg Val Gly Arg Arg Asp Asp His Gly Gly Tyr Arg Arg Ala
        35                  40                  45
Gln Phe Asp His Val Leu Ser Ser Pro Gly Thr His Gly Asp Asp Gln
    50                  55                  60
Val Asp Val Leu Gly Ala Met Gln Ala Cys Arg Asn Gly Phe Ala His
65                  70                  75                  80
Gly Arg Pro Gln Gly Gly Val Glu Val Ala Gly Leu Pro Pro Arg Ile
                85                  90                  95
Gly His Glu Leu Asp Val Val Ala Ala Gly Met Arg Gly Gln Val Lys
            100                 105                 110
Pro Gly Leu Ser His Ile Gly Tyr Leu Gly His Val Val Asp Thr Gln
        115                 120                 125
Pro Ser Glu Leu Leu Gln Cys Ala Arg Asp Arg Val Glu Arg Ala Arg
    130                 135                 140
Ile Leu Leu Val Leu Ala Ala Thr His Ile Thr Asp Ala Pro Phe Pro
145                 150                 155                 160
Ala Gly Cys Met Gly Gly Arg Arg Lys Ala Gly Arg Gly Lys Gly Asp
                165                 170                 175
Ala Gly Ala Gln Gly Arg Asp Arg Arg Arg Ala Gly Cys Arg Arg Asp
            180                 185                 190
Gly Pro Arg Glu Ala Lys Lys Gln Gly Glu Arg Gly Glu Tyr Ser Gly
```

```
          195                 200                 205
Ala Lys Gly His Gly Arg
    210

      <210> 73
      <211> 1314
      <212> DNA
      <213> Bordetella pertussis

      <220>
      <221> CDS
      <222> (1)...(1314)

      <400> 73
atg tcc aat acc tat ttc ccg cgc tgg cgg ctg gcc gac gac acc gtg      48
Met Ser Asn Thr Tyr Phe Pro Arg Trp Arg Leu Ala Asp Asp Thr Val
 1               5                  10                  15

ccg ggc gcg gtc atc gcg ccc gac gaa cgc ctg tcc tgg ccc aag aac      96
Pro Gly Ala Val Ile Ala Pro Asp Glu Arg Leu Ser Trp Pro Lys Asn
            20                  25                  30

atc gcc atg ggg gcc cag cac gtg gtc gcc atg ttc ggt tcc acc gtg     144
Ile Ala Met Gly Ala Gln His Val Val Ala Met Phe Gly Ser Thr Val
        35                  40                  45

ctg gcg ccg ctg ctg atg ggt ttc gac ccc aat gtg gcg atc ctc atg     192
Leu Ala Pro Leu Leu Met Gly Phe Asp Pro Asn Val Ala Ile Leu Met
    50                  55                  60

tcc ggc atc ggc acg ctg atc ttc ttc ctg ttc gtc ggc ggc cgg gtg     240
Ser Gly Ile Gly Thr Leu Ile Phe Phe Leu Phe Val Gly Gly Arg Val
65                  70                  75                  80

ccc agc tac ctg ggc tcc agc ttc gcc ttc atc ggc ggg gtg gtg gcg     288
Pro Ser Tyr Leu Gly Ser Ser Phe Ala Phe Ile Gly Gly Val Val Ala
                85                  90                  95

gtc acc ggc tat gtg gcg ccc ggc gcc aac gcc aat atc ggc gtg gcg     336
Val Thr Gly Tyr Val Ala Pro Gly Ala Asn Ala Asn Ile Gly Val Ala
            100                 105                 110

ctc ggc gcg atc atc gcc tgt ggc ctg gtg tac gcg ctg atc ggc ctg     384
Leu Gly Ala Ile Ile Ala Cys Gly Leu Val Tyr Ala Leu Ile Gly Leu
        115                 120                 125

gtc gta tgg gcg gcc agc gcg cgc ggc aac ggg gcg cgc tgg atc gag     432
Val Val Trp Ala Ala Ser Ala Arg Gly Asn Gly Ala Arg Trp Ile Glu
    130                 135                 140

gcc atg atg ccg ccg gtc gtc acg ggc gcg gtg gtg gcg gtg atc ggc     480
Ala Met Met Pro Pro Val Val Thr Gly Ala Val Val Ala Val Ile Gly
145                 150                 155                 160

ctg aac ctg gcc ccg atc gcc gcc aag ggc gcc atg ggt tcg tcc ggc     528
Leu Asn Leu Ala Pro Ile Ala Ala Lys Gly Ala Met Gly Ser Ser Gly
                165                 170                 175

ttc gag gcc agc atg gcg ttg atg acc atc ctg tgc gtg ggc ggc atc     576
Phe Glu Ala Ser Met ala Leu Met Thr Ile Leu Cys Val Gly Gly Ile
```

```
                180                 185                 190
gcc gtc tac acg cgc ggc atg gtg cag cgg ctg ctg atc ctg gtc ggc      624
Ala Val Tyr Thr Arg Gly Met Val Gln Arg Leu Leu Ile Leu Val Gly
        195                 200                 205

ctg gtg ctg gcc tgc gtc atc tac gcg gtc tgc gcc aac ggc ctg ggg      672
Leu Val Leu Ala Cys Val Ile Tyr Ala Val Cys Ala Asn Gly Leu Gly
    210                 215                 220

ctg ggc gcg ccc atg gac ttc gcc aag gtg gcc gcc gcg ccg tgg ttc      720
Leu Gly Ala Pro Met Asp Phe Ala Lys Val Ala Ala Ala Pro Trp Phe
225                 230                 235                 240

ggc ctg ccc agc ttc gcc gcg ccg gtg ttc gag ccg cag gcc atg ggc      768
Gly Leu Pro Ser Phe Ala Ala Pro Val Phe Glu Pro Gln Ala Met Gly
                245                 250                 255

ctg atc gtg ccg gtg gcc atc atc ctg gtg gcc gag aac ctg ggc cac      816
Leu Ile Val Pro Val Ala Ile Ile Leu Val Ala Glu Asn Leu Gly His
            260                 265                 270

gtg aag gcg gtc gcc gcc atg acc gga cag gac ctg gac cgc tac gtg      864
Val Lys Ala Val Ala Ala Met Thr Gly Gln Asp Leu Asp Arg Tyr Val
        275                 280                 285

ggc cgc gcc ttc gtg ggc gac ggc gtg gcg acc atg gtt tcc ggc gcc      912
Gly Arg Ala Phe Val Gly Asp Gly Val Ala Thr Met Val Ser Gly Ala
    290                 295                 300

gtc ggc ggc acc ggg gtg acc acc tac gcc gag aat atc ggc gtg atg      960
Val Gly Gly Thr Gly Val Thr Thr Tyr Ala Glu Asn Ile Gly Val Met
305                 310                 315                 320

gcc gtg acg cgc atc tat tcc acg ctg gtg ttc gtg gtg gcg gcc gtg     1008
Ala Val Thr Arg Ile Tyr Ser Thr Leu Val Phe Val Val Ala Ala Val
                325                 330                 335

atc gcg ctg gtg ctg ggg ttc tcg ccc aag ttc ggc gcg ctg atc cag     1056
Ile Ala Leu Val Leu Gly Phe Ser Pro Lys Phe Gly Ala Leu Ile Gln
            340                 345                 350

acc atc ccc ggc ccc gtg ctg ggg ggc atg tcg gtc gtg gtg ttc ggc     1104
Thr Ile Pro Gly Pro Val Leu Gly Gly Met Ser Val Val Val Phe Gly
        355                 360                 365

ctg atc gcc atc gcc ggc gcg cgc atc tgg gtg gtc aac cag gtc gat     1152
Leu Ile Ala Ile Ala Gly Ala Arg Ile Trp Val Val Asn Gln Val Asp
    370                 375                 380

ttc agc gac aac cgc aat ctg atc gtg gcc gcc gtg acc ctg gtg ctg     1200
Phe Ser Asp Asn Arg Asn Leu Ile Val Ala Ala Val Thr Leu Val Leu
385                 390                 395                 400

ggg gcg ggc gac ttc agc gtc aag ctg ggc gat ttc tcg atg aac ggc     1248
Gly Ala Gly Asp Phe Ser Val Lys Leu Gly Asp Phe Ser Met Asn Gly
                405                 410                 415

atc ggc acc gcc acg ttc ggc gcc atc atc ctg tac gcc ctg ctg ggc     1296
Ile Gly Thr Ala Thr Phe Gly Ala Ile Ile Leu Tyr Ala Leu Leu Gly
```

420 425 430

ctg gcg cgt cgc cgc tga 1314
Leu Ala Arg Arg Arg *
435

<210> 74
<211> 437
<212> PRT
<213> Bordetella pertussis

<400> 74

Met	Ser	Asn	Thr	Tyr	Phe	Pro	Arg	Trp	Arg	Leu	Ala	Asp	Asp	Thr	Val
1				5				10						15	
Pro	Gly	Ala	Val	Ile	Ala	Pro	Asp	Glu	Arg	Leu	Ser	Trp	Pro	Lys	Asn
			20					25					30		
Ile	Ala	Met	Gly	Ala	Gln	His	Val	Val	Ala	Met	Phe	Gly	Ser	Thr	Val
		35					40					45			
Leu	Ala	Pro	Leu	Leu	Met	Gly	Phe	Asp	Pro	Asn	Val	Ala	Ile	Leu	Met
	50					55					60				
Ser	Gly	Ile	Gly	Thr	Leu	Ile	Phe	Phe	Leu	Phe	Val	Gly	Gly	Arg	Val
65					70					75					80
Pro	Ser	Tyr	Leu	Gly	Ser	Ser	Phe	Ala	Phe	Ile	Gly	Gly	Val	Val	Ala
				85					90					95	
Val	Thr	Gly	Tyr	Val	Ala	Pro	Gly	Ala	Asn	Ala	Asn	Ile	Gly	Val	Ala
			100					105					110		
Leu	Gly	Ala	Ile	Ile	Ala	Cys	Gly	Leu	Val	Tyr	Ala	Leu	Ile	Gly	Leu
		115					120					125			
Val	Val	Trp	Ala	Ala	Ser	Ala	Arg	Gly	Asn	Gly	Ala	Arg	Trp	Ile	Glu
	130					135					140				
Ala	Met	Met	Pro	Pro	Val	Val	Thr	Gly	Ala	Val	Val	Ala	Val	Ile	Gly
145					150					155					160
Leu	Asn	Leu	Ala	Pro	Ile	Ala	Ala	Lys	Gly	Ala	Met	Gly	Ser	Ser	Gly
				165					170					175	
Phe	Glu	Ala	Ser	Met	ala	Leu	Met	Thr	Ile	Leu	Cys	Val	Gly	Gly	Ile
			180					185					190		
Ala	Val	Tyr	Thr	Arg	Gly	Met	Val	Gln	Arg	Leu	Leu	Ile	Leu	Val	Gly
		195					200					205			
Leu	Val	Leu	Ala	Cys	Val	Ile	Tyr	Ala	Val	Cys	Ala	Asn	Gly	Leu	Gly
	210					215					220				
Leu	Gly	Ala	Pro	Met	Asp	Phe	Ala	Lys	Val	Ala	Ala	Ala	Pro	Trp	Phe
225					230					235					240
Gly	Leu	Pro	Ser	Phe	Ala	Ala	Pro	Val	Phe	Glu	Pro	Gln	Ala	Met	Gly
				245					250					255	
Leu	Ile	Val	Pro	Val	Ala	Ile	Ile	Leu	Val	Ala	Glu	Asn	Leu	Gly	His
			260					265					270		
Val	Lys	Ala	Val	Ala	Ala	Met	Thr	Gly	Gln	Asp	Leu	Asp	Arg	Tyr	Val
		275					280					285			
Gly	Arg	Ala	Phe	Val	Gly	Asp	Gly	Val	Ala	Thr	Met	Val	Ser	Gly	Ala
	290					295					300				
Val	Gly	Gly	Thr	Gly	Val	Thr	Thr	Tyr	Ala	Glu	Asn	Ile	Gly	Val	Met
305					310					315					320
Ala	Val	Thr	Arg	Ile	Tyr	Ser	Thr	Leu	Val	Phe	Val	Val	Ala	Ala	Val
				325					330					335	
Ile	Ala	Leu	Val	Leu	Gly	Phe	Ser	Pro	Lys	Phe	Gly	Ala	Leu	Ile	Gln
			340					345					350		
Thr	Ile	Pro	Gly	Pro	Val	Leu	Gly	Gly	Met	Ser	Val	Val	Val	Phe	Gly
		355					360					365			
Leu	Ile	Ala	Ile	Ala	Gly	Ala	Arg	Ile	Trp	Val	Val	Asn	Gln	Val	Asp

```
    370                 375                 380
Phe Ser Asp Asn Arg Asn Leu Ile Val Ala Ala Val Thr Leu Val Leu
385                 390                 395                 400
Gly Ala Gly Asp Phe Ser Val Lys Leu Gly Asp Phe Ser Met Asn Gly
                405                 410                 415
Ile Gly Thr Ala Thr Phe Gly Ala Ile Ile Leu Tyr Ala Leu Leu Gly
            420                 425                 430
Leu Ala Arg Arg Arg
        435
```

<210> 75
<211> 1536
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(1536)

<400> 75

```
atg ctt gaa agg atc aag gtc cgc acc gcg atg gtg gcg gta ttc gcg        48
Met Leu Glu Arg Ile Lys Val Arg Thr Ala Met Val Ala Val Phe Ala
 1               5                  10                  15

tgc ttc ctg gcg gtg ctg atg ctg tcg ggc gcc ctg acg tgg cgc aac        96
Cys Phe Leu Ala Val Leu Met Leu Ser Gly Ala Leu Thr Trp Arg Asn
            20                  25                  30

gcg ggc agg agc gcc gcc gag atc gag ggg ctg aac cag gtc gcc gtc       144
Ala Gly Arg Ser Ala Ala Glu Ile Glu Gly Leu Asn Gln Val Ala Val
        35                  40                  45

aac cag gtc gac ccg ctg ttc gag gcc agc ggc gcg gcg cag cgc cag       192
Asn Gln Val Asp Pro Leu Phe Glu Ala Ser Gly Ala Ala Gln Arg Gln
    50                  55                  60

gcg gcc acg caa ttc cag cgc tac gtg gac gtg ccc aag gag ccg gcc       240
Ala Ala Thr Gln Phe Gln Arg Tyr Val Asp Val Pro Lys Glu Pro Ala
65                  70                  75                  80

gcg gcc gag ctg gcc gcg acc ctg cag acg cgc tgg cgc gcc tac cag       288
Ala Ala Glu Leu Ala Ala Thr Leu Gln Thr Arg Trp Arg Ala Tyr Gln
                85                  90                  95

tcg gtg ctg gac gag ctg gcc gcc gcc gtc gac gcc ggc cag gcc gag       336
Ser Val Leu Asp Glu Leu Ala Ala Ala Val Asp Ala Gly Gln Ala Glu
            100                 105                 110

ccc gcc ctg gcc gcc atg cat cgc gcg cag cag gcc gaa cat gca ttc       384
Pro Ala Leu Ala Ala Met His Arg Ala Gln Gln Ala Glu His Ala Phe
        115                 120                 125

cag cgc gac atg gaa gcc ttt ctg gcc agg gta cag gcg cac agc gac       432
Gln Arg Asp Met Glu Ala Phe Leu Ala Arg Val Gln Ala His Ser Asp
    130                 135                 140

gaa gtg cgc agc ggc gcc gag gac acc cat gtc gtg gcc cgc tgg agc       480
Glu Val Arg Ser Gly Ala Glu Asp Thr His Val Val Ala Arg Trp Ser
145                 150                 155                 160
```

```
gcc atc gcg ctg acc acg ctg ggc gtg ctg ctg acc ctg gcc ggc tgg      528
Ala Ile Ala Leu Thr Thr Leu Gly Val Leu Leu Thr Leu Ala Gly Trp
                165                 170                 175

ctg ttc gtg cgc cgc gcg gtg ctg cgc ccc ttg ctg gag gcc ggc cat      576
Leu Phe Val Arg Arg Ala Val Leu Arg Pro Leu Leu Glu Ala Gly His
            180                 185                 190

cat ttc gac cgc atc gcc gac ggc gac ctc acc gcg cgc atc gag gtg      624
His Phe Asp Arg Ile Ala Asp Gly Asp Leu Thr Ala Arg Ile Glu Val
        195                 200                 205

cgc tcg gcc aat gaa atc ggc gcg ctg ttc gcg gcg ctc aag cgc atg      672
Arg Ser Ala Asn Glu Ile Gly Ala Leu Phe Ala Ala Leu Lys Arg Met
    210                 215                 220

cag gaa ggc ctg acg cgc acc atc gcc gtc atg cgg cgc ggc gtc gac      720
Gln Glu Gly Leu Thr Arg Thr Ile Ala Val Met Arg Arg Gly Val Asp
225                 230                 235                 240

gaa atc aac gtc ggc gcg gcc gag atc tcg gcc ggc aac gcc aac ctg      768
Glu Ile Asn Val Gly Ala Ala Glu Ile Ser Ala Gly Asn Ala Asn Leu
                245                 250                 255

tcc agc cgc acg gag gag cag gcc gcc gcc ctg gaa gag acc gcg gcc      816
Ser Ser Arg Thr Glu Glu Gln Ala Ala Ala Leu Glu Glu Thr Ala Ala
            260                 265                 270

acc atg gag gaa ctg gcc acc acg gtc aag cag aac gcc gac aat gcc      864
Thr Met Glu Glu Leu Ala Thr Thr Val Lys Gln Asn Ala Asp Asn Ala
        275                 280                 285

gcg cag gcc aat cag ctg gcc gcc gtc agc atg cag gtg gcg cag cgc      912
Ala Gln Ala Asn Gln Leu Ala Ala Val Ser Met Gln Val Ala Gln Arg
    290                 295                 300

ggc ggc gag tcg gtc gcg cag gtg gtg cag acc atg cac ggc atc tcc      960
Gly Gly Glu Ser Val Ala Gln Val Val Gln Thr Met His Gly Ile Ser
305                 310                 315                 320

gcg agc tcg cgc cag atc gcc gac atc gtc acc gtg atc gac ggc atc     1008
Ala Ser Ser Arg Gln Ile Ala Asp Ile Val Thr Val Ile Asp Gly Ile
                325                 330                 335

gcc ttc cag acc aat atc ctg gcg ctg aac gcc gcg gtc gag gcg gcg     1056
Ala Phe Gln Thr Asn Ile Leu Ala Leu Asn Ala Ala Val Glu Ala Ala
            340                 345                 350

cgc gcc ggc gaa cag ggc aag ggc ttc gcg gtg gtg gcg ggc gag gtg     1104
Arg Ala Gly Glu Gln Gly Lys Gly Phe Ala Val Val Ala Gly Glu Val
        355                 360                 365

cgc agc ctg gcc cag cgc gcc gcg cag gcg gcc aag gag atc aag gcc     1152
Arg Ser Leu Ala Gln Arg Ala Ala Gln Ala Ala Lys Glu Ile Lys Ala
    370                 375                 380

ctg atc gag agc tcg gtg gcg acg gtg cgc gcc ggc tcg caa cag gtc     1200
Leu Ile Glu Ser Ser Val Ala Thr Val Arg Ala Gly Ser Gln Gln Val
385                 390                 395                 400
```

```
gcc agc gcc ggc ggc acc atg gac gag gtg gtg gcc tcg gta cag cgc      1248
Ala Ser Ala Gly Gly Thr Met Asp Glu Val Val Ala Ser Val Gln Arg
                405                 410                 415

gtg gcc gac atc atg ggg gag atc tcg gcc gcc tcg gcc cag cag gcc      1296
Val Ala Asp Ile Met Gly Glu Ile Ser Ala Ala Ser Ala Gln Gln Ala
            420                 425                 430

agc ggc atc gac cag gtc agc ctg gcg att tcg caa atg gac gaa acc      1344
Ser Gly Ile Asp Gln Val Ser Leu Ala Ile Ser Gln Met Asp Glu Thr
        435                 440                 445

acc cag cag aat gcc gcg ctg gtc gaa cag gcc gcg gcg gcg gcc acg      1392
Thr Gln Gln Asn Ala Ala Leu Val Glu Gln Ala Ala Ala Ala Ala Thr
    450                 455                 460

gcc atg gaa gaa cag gcc cgc cac ctg gcg gcc gcg gcg gcg gtc ttc      1440
Ala Met Glu Glu Gln Ala Arg His Leu Ala Ala Ala Ala Ala Val Phe
465                 470                 475                 480

agg acg cag ggc ggc gcc atc atc gac gtc gcc gcc gcg ccg ctg gcc      1488
Arg Thr Gln Gly Gly Ala Ile Ile Asp Val Ala Ala Ala Pro Leu Ala
                485                 490                 495

ggg ccg gcg ggc ggc cat gcc gcc ctg ccg ccg gcc gcg gcc cac tga      1536
Gly Pro Ala Gly Gly His Ala Ala Leu Pro Pro Ala Ala Ala His  *
            500                 505                 510
```

<210> 76
<211> 511
<212> PRT
<213> Bordetella pertussis

<400> 76

```
Met Leu Glu Arg Ile Lys Val Arg Thr Ala Met Val Ala Val Phe Ala
 1               5                  10                  15
Cys Phe Leu Ala Val Leu Met Leu Ser Gly Ala Leu Thr Trp Arg Asn
            20                  25                  30
Ala Gly Arg Ser Ala Ala Glu Ile Glu Gly Leu Asn Gln Val Ala Val
        35                  40                  45
Asn Gln Val Asp Pro Leu Phe Glu Ala Ser Gly Ala Ala Gln Arg Gln
    50                  55                  60
Ala Ala Thr Gln Phe Gln Arg Tyr Val Asp Val Pro Lys Glu Pro Ala
65                  70                  75                  80
Ala Ala Glu Leu Ala Ala Thr Leu Gln Thr Arg Trp Arg Ala Tyr Gln
                85                  90                  95
Ser Val Leu Asp Glu Leu Ala Ala Ala Val Asp Ala Gly Gln Ala Glu
            100                 105                 110
Pro Ala Leu Ala Ala Met His Arg Ala Gln Gln Ala Glu His Ala Phe
        115                 120                 125
Gln Arg Asp Met Glu Ala Phe Leu Ala Arg Val Gln Ala His Ser Asp
    130                 135                 140
Glu Val Arg Ser Gly Ala Glu Asp Thr His Val Val Ala Arg Trp Ser
145                 150                 155                 160
Ala Ile Ala Leu Thr Thr Leu Gly Val Leu Leu Thr Leu Ala Gly Trp
                165                 170                 175
Leu Phe Val Arg Arg Ala Val Leu Arg Pro Leu Leu Glu Ala Gly His
            180                 185                 190
His Phe Asp Arg Ile Ala Asp Gly Asp Leu Thr Ala Arg Ile Glu Val
```

```
            195                 200                 205
Arg Ser Ala Asn Glu Ile Gly Ala Leu Phe Ala Ala Leu Lys Arg Met
        210                 215                 220
Gln Glu Gly Leu Thr Arg Thr Ile Ala Val Met Arg Arg Gly Val Asp
225                 230                 235                 240
Glu Ile Asn Val Gly Ala Ala Glu Ile Ser Ala Gly Asn Ala Asn Leu
                245                 250                 255
Ser Ser Arg Thr Glu Glu Gln Ala Ala Ala Leu Glu Glu Thr Ala Ala
            260                 265                 270
Thr Met Glu Glu Leu Ala Thr Thr Val Lys Gln Asn Ala Asp Asn Ala
        275                 280                 285
Ala Gln Ala Asn Gln Leu Ala Ala Val Ser Met Gln Val Ala Gln Arg
    290                 295                 300
Gly Gly Glu Ser Val Ala Gln Val Val Gln Thr Met His Gly Ile Ser
305                 310                 315                 320
Ala Ser Ser Arg Gln Ile Ala Asp Ile Val Thr Val Ile Asp Gly Ile
                325                 330                 335
Ala Phe Gln Thr Asn Ile Leu Ala Leu Asn Ala Ala Val Glu Ala Ala
            340                 345                 350
Arg Ala Gly Glu Gln Gly Lys Gly Phe Ala Val Val Ala Gly Glu Val
        355                 360                 365
Arg Ser Leu Ala Gln Arg Ala Ala Gln Ala Ala Lys Glu Ile Lys Ala
    370                 375                 380
Leu Ile Glu Ser Ser Val Ala Thr Val Arg Ala Gly Ser Gln Gln Val
385                 390                 395                 400
Ala Ser Ala Gly Gly Thr Met Asp Glu Val Val Ala Ser Val Gln Arg
                405                 410                 415
Val Ala Asp Ile Met Gly Glu Ile Ser Ala Ala Ser Ala Gln Gln Ala
            420                 425                 430
Ser Gly Ile Asp Gln Val Ser Leu Ala Ile Ser Gln Met Asp Glu Thr
        435                 440                 445
Thr Gln Gln Asn Ala Ala Leu Val Glu Gln Ala Ala Ala Ala Ala Thr
    450                 455                 460
Ala Met Glu Glu Gln Ala Arg His Leu Ala Ala Ala Ala Ala Val Phe
465                 470                 475                 480
Arg Thr Gln Gly Gly Ala Ile Ile Asp Val Ala Ala Ala Pro Leu Ala
                485                 490                 495
Gly Pro Ala Gly Gly His Ala Ala Leu Pro Pro Ala Ala Ala His
            500                 505                 510
```

<210> 77
<211> 477
<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(477)

<400> 77

```
atg tct gcc att cct ttg acc gtg cgc ggg gcc gag cgc ttg cag caa      48
Met Ser Ala Ile Pro Leu Thr Val Arg Gly Ala Glu Arg Leu Gln Gln
 1               5                  10                  15

gaa ctg cat cgg ctt aag acc gtt gag cgt cct gcg gtg atc agc gcc      96
Glu Leu His Arg Leu Lys Thr Val Glu Arg Pro Ala Val Ile Ser Ala
            20                  25                  30

att gcg gag gcg cgt gcg cag ggt gat ttg tcg gaa aat gcc gag tac     144
Ile Ala Glu Ala Arg Ala Gln Gly Asp Leu Ser Glu Asn Ala Glu Tyr
```

```
                35                  40                  45

gac gcc gcc cgc gaa cgc cag ggc ttc atc gaa ggc cgg atc tcc gaa      192
Asp Ala Ala Arg Glu Arg Gln Gly Phe Ile Glu Gly Arg Ile Ser Glu
     50                  55                  60

ctc gag ggc acg ctt tcg aac gcg cac ctc atc gat cca acg gcg ctc      240
Leu Glu Gly Thr Leu Ser Asn Ala His Leu Ile Asp Pro Thr Ala Leu
 65                  70                  75                  80

gac gcc gaa ggc cgt gcc gtg ttc ggc gcg acc gtg gaa atc gaa gac      288
Asp Ala Glu Gly Arg Ala Val Phe Gly Ala Thr Val Glu Ile Glu Asp
                 85                  90                  95

ctc gac tcg ggc gac cgc ctg acc tac cag atc gtg ggc gac gtc gaa      336
Leu Asp Ser Gly Asp Arg Leu Thr Tyr Gln Ile Val Gly Asp Val Glu
            100                 105                 110

gcc gac atc aag tcc aac ctg att tcg gtc tcc agc ccg gtg gcc cgc      384
Ala Asp Ile Lys Ser Asn Leu Ile Ser Val Ser Ser Pro Val Ala Arg
        115                 120                 125

gcc ctg atc ggc aaa tcc gag ggc gat gtg gtc gaa gtg aag gtg ccg      432
Ala Leu Ile Gly Lys Ser Glu Gly Asp Val Val Glu Val Lys Val Pro
    130                 135                 140

gct ggc gtg cgc gag tac gaa gtc atc ggt gtg cgt tat ctc tga          477
Ala Gly Val Arg Glu Tyr Glu Val Ile Gly Val Arg Tyr Leu  *
145                 150                 155
```

<210> 78
<211> 158
<212> PRT
<213> Bordetella pertussis

<400> 78

```
Met Ser Ala Ile Pro Leu Thr Val Arg Gly Ala Glu Arg Leu Gln Gln
 1               5                  10                  15
Glu Leu His Arg Leu Lys Thr Val Glu Arg Pro Ala Val Ile Ser Ala
            20                  25                  30
Ile Ala Glu Ala Arg Ala Gln Gly Asp Leu Ser Glu Asn Ala Glu Tyr
        35                  40                  45
Asp Ala Ala Arg Glu Arg Gln Gly Phe Ile Glu Gly Arg Ile Ser Glu
    50                  55                  60
Leu Glu Gly Thr Leu Ser Asn Ala His Leu Ile Asp Pro Thr Ala Leu
65                  70                  75                  80
Asp Ala Glu Gly Arg Ala Val Phe Gly Ala Thr Val Glu Ile Glu Asp
                85                  90                  95
Leu Asp Ser Gly Asp Arg Leu Thr Tyr Gln Ile Val Gly Asp Val Glu
            100                 105                 110
Ala Asp Ile Lys Ser Asn Leu Ile Ser Val Ser Ser Pro Val Ala Arg
        115                 120                 125
Ala Leu Ile Gly Lys Ser Glu Gly Asp Val Val Glu Val Lys Val Pro
    130                 135                 140
Ala Gly Val Arg Glu Tyr Glu Val Ile Gly Val Arg Tyr Leu
145                 150                 155
```

<210> 79
<211> 951

<212> DNA
<213> Bordetella pertussis

<220>
<221> CDS
<222> (1)...(951)

<400> 79

```
atg aac acc cat aag cat gcc cga ttg acc ttc cta cgt cga ctc gaa        48
Met Asn Thr His Lys His Ala Arg Leu Thr Phe Leu Arg Arg Leu Glu
 1               5                  10                  15

atg gtc cag caa ttg atc gcc cat caa gtt tgt gtg cct gaa gcg gcc        96
Met Val Gln Gln Leu Ile Ala His Gln Val Cys Val Pro Glu Ala Ala
            20                  25                  30

cgc gcc tat ggg gtc acc gcg ccg act gtg cgc aaa tgg ctg ggc cgc       144
Arg Ala Tyr Gly Val Thr Ala Pro Thr Val Arg Lys Trp Leu Gly Arg
        35                  40                  45

ttc ctg gct cag ggc cag gcg ggc ttg gcc gat gcg tcc tcg cgc ccg       192
Phe Leu Ala Gln Gly Gln Ala Gly Leu Ala Asp Ala Ser Ser Arg Pro
    50                  55                  60

acg gtc tcg ccc cga gcg att gcg ccg gcc aag gcg ctg gct atc gtg       240
Thr Val Ser Pro Arg Ala Ile Ala Pro Ala Lys Ala Leu Ala Ile Val
65                  70                  75                  80

gag ctg cgc cgc aag cgg ctg acc caa gcg cgc atc gcc cag gcg ctg       288
Glu Leu Arg Arg Lys Arg Leu Thr Gln Ala Arg Ile Ala Gln Ala Leu
                85                  90                  95

ggc gtg tca gcc agc acc gtc agc cgc gtc ctg gcc cgc gcc ggt ctg       336
Gly Val Ser Ala Ser Thr Val Ser Arg Val Leu Ala Arg Ala Gly Leu
            100                 105                 110

tcg cac ctg gcc gac ctg gag ccg gcc gag ccg gtg gtg cgc tac gag       384
Ser His Leu Ala Asp Leu Glu Pro Ala Glu Pro Val Val Arg Tyr Glu
        115                 120                 125

cat cag gcc ccc ggc gat ctg ctg cac atc gac atc aag aag ctg gga       432
His Gln Ala Pro Gly Asp Leu Leu His Ile Asp Ile Lys Lys Leu Gly
    130                 135                 140

cgt atc cag cgc cct ggc cac cgg gtc acg ggc aac cga cgc gat acc       480
Arg Ile Gln Arg Pro Gly His Arg Val Thr Gly Asn Arg Arg Asp Thr
145                 150                 155                 160

gtt gag ggg gcc ggc tgg gac ttc gtc ttc gtg gcc atc gat gac cac       528
Val Glu Gly Ala Gly Trp Asp Phe Val Phe Val Ala Ile Asp Asp His
                165                 170                 175

gcc cgc gtg gcc ttc acc gac atc cac ccc gac gag cgc ttc ccc agc       576
Ala Arg Val Ala Phe Thr Asp Ile His Pro Asp Glu Arg Phe Pro Ser
            180                 185                 190

gcc gtc cag ttc ctc aag gac gca gtg gcc tac tac cag cgc ctg ggc       624
Ala Val Gln Phe Leu Lys Asp Ala Val Ala Tyr Tyr Gln Arg Leu Gly
        195                 200                 205
```

```
gtg acc atc cag cgc ttg ctc acc gac aat ggc tcg gcc ttt cgc agc      672
Val Thr Ile Gln Arg Leu Leu Thr Asp Asn Gly Ser Ala Phe Arg Ser
    210                 215                 220

cgc gcc ttc gcc gcg ctg tgc cat gag ctg ggc atc aag cac cgc ttt      720
Arg Ala Phe Ala Ala Leu Cys His Glu Leu Gly Ile Lys His Arg Phe
225                 230                 235                 240

acc cga cct tac cgc cca cag acc aat ggc aag gcc gaa cgc ttc atc      768
Thr Arg Pro Tyr Arg Pro Gln Thr Asn Gly Lys Ala Glu Arg Phe Ile
                245                 250                 255

cag tcg gcc ttg cgt gag tgg gct tac gct cac acc tac cag aac tcc      816
Gln Ser Ala Leu Arg Glu Trp Ala Tyr Ala His Thr Tyr Gln Asn Ser
            260                 265                 270

caa cac cga gcc gat gcc atg aaa tcc tgg cta cac cac tac aac tgg      864
Gln His Arg Ala Asp Ala Met Lys Ser Trp Leu His His Tyr Asn Trp
        275                 280                 285

cat cga ccc cac caa ggc atc ggg cgc gct gta ccc atc tcc aga ctc      912
His Arg Pro His Gln Gly Ile Gly Arg Ala Val Pro Ile Ser Arg Leu
    290                 295                 300

aac ctg gac gaa tac aac cta ttg aca gtt cac acc tag                  951
Asn Leu Asp Glu Tyr Asn Leu Leu Thr Val His Thr  *
305                 310                 315
```

<210> 80
<211> 316
<212> PRT
<213> Bordetella pertussis

<400> 80

```
Met Asn Thr His Lys His Ala Arg Leu Thr Phe Leu Arg Arg Leu Glu
 1               5                  10                  15
Met Val Gln Gln Leu Ile Ala His Gln Val Cys Val Pro Glu Ala Ala
            20                  25                  30
Arg Ala Tyr Gly Val Thr Ala Pro Thr Val Arg Lys Trp Leu Gly Arg
        35                  40                  45
Phe Leu Ala Gln Gly Gln Ala Gly Leu Ala Asp Ala Ser Ser Arg Pro
    50                  55                  60
Thr Val Ser Pro Arg Ala Ile Ala Pro Ala Lys Ala Leu Ala Ile Val
65                  70                  75                  80
Glu Leu Arg Arg Lys Arg Leu Thr Gln Ala Arg Ile Ala Gln Ala Leu
                85                  90                  95
Gly Val Ser Ala Ser Thr Val Ser Arg Val Leu Ala Arg Ala Gly Leu
            100                 105                 110
Ser His Leu Ala Asp Leu Glu Pro Ala Glu Pro Val Val Arg Tyr Glu
        115                 120                 125
His Gln Ala Pro Gly Asp Leu Leu His Ile Asp Ile Lys Lys Leu Gly
    130                 135                 140
Arg Ile Gln Arg Pro Gly His Arg Val Thr Gly Asn Arg Arg Asp Thr
145                 150                 155                 160
Val Glu Gly Ala Gly Trp Asp Phe Val Phe Val Ala Ile Asp Asp His
                165                 170                 175
Ala Arg Val Ala Phe Thr Asp Ile His Pro Asp Glu Arg Phe Pro Ser
            180                 185                 190
Ala Val Gln Phe Leu Lys Asp Ala Val Ala Tyr Tyr Gln Arg Leu Gly
```

```
          195                 200                 205
Val Thr Ile Gln Arg Leu Leu Thr Asp Asn Gly Ser Ala Phe Arg Ser
    210                 215                 220
Arg Ala Phe Ala Ala Leu Cys His Glu Leu Gly Ile Lys His Arg Phe
225                 230                 235                 240
Thr Arg Pro Tyr Arg Pro Gln Thr Asn Gly Lys Ala Glu Arg Phe Ile
                245                 250                 255
Gln Ser Ala Leu Arg Glu Trp Ala Tyr Ala His Thr Tyr Gln Asn Ser
            260                 265                 270
Gln His Arg Ala Asp Ala Met Lys Ser Trp Leu His His Tyr Asn Trp
        275                 280                 285
His Arg Pro His Gln Gly Ile Gly Arg Ala Val Pro Ile Ser Arg Leu
    290                 295                 300
Asn Leu Asp Glu Tyr Asn Leu Leu Thr Val His Thr
305                 310                 315

      <210> 81
      <211> 1851
      <212> DNA
      <213> Bordetella pertussis

      <220>
      <221> CDS
      <222> (1)...(1851)

      <400> 81
atg gac ttg gta gtt cgt gac acc gac acg cgc tgg tcc acg ctg ctc      48
Met Asp Leu Val Val Arg Asp Thr Asp Thr Arg Trp Ser Thr Leu Leu
 1               5                  10                  15

gac gac aag atc cgc acc atc cgg gaa agc cgc agg caa ctg atc caa      96
Asp Asp Lys Ile Arg Thr Ile Arg Glu Ser Arg Arg Gln Leu Ile Gln
            20                  25                  30

ctc agc gcg gtc gtg aca tcg gtg ctg aac gcc tat gcc gcg cag gcc     144
Leu Ser Ala Val Val Thr Ser Val Leu Asn Ala Tyr Ala Ala Gln Ala
        35                  40                  45

gag cgc gga cac gtc act acc ggc gcc gcc aag ggc atg gcg cgt gtc     192
Glu Arg Gly His Val Thr Thr Gly Ala Ala Lys Gly Met ala Arg Val
    50                  55                  60

tgg ctg aac cat ctc gac ctg gga ccg cgc cgc gtc gcc ttc gcc tat     240
Trp Leu Asn His Leu Asp Leu Gly Pro Arg Arg Val Ala Phe Ala Tyr
65                  70                  75                  80

gac gcg gaa ggc acc gtg ctg gcc agc acc aac ccc cgg atg atc gac     288
Asp Ala Glu Gly Thr Val Leu Ala Ser Thr Asn Pro Arg Met Ile Asp
                85                  90                  95

cgg gac ctc tcc ggg atc cgc gac ttc aag ggc cgg ccg ctc gcc gcc     336
Arg Asp Leu Ser Gly Ile Arg Asp Phe Lys Gly Arg Pro Leu Ala Ala
            100                 105                 110

gcc atg tac gag gaa agc cgc aac gac ggt cgc ggc ttc gcc atc tac     384
Ala Met Tyr Glu Glu Ser Arg Asn Asp Gly Arg Gly Phe Ala Ile Tyr
        115                 120                 125

ccg tct ccc ctg gac gag tcc gcc cag atg cga cac gcc tat ttc gtg     432
Pro Ser Pro Leu Asp Glu Ser Ala Gln Met Arg His Ala Tyr Phe Val
```

```
     130                 135                 140

tac ttc ccc gcg tgg aag tgg gtt ctc gcc atc tcc gat agc tca caa      480
Tyr Phe Pro Ala Trp Lys Trp Val Leu Ala Ile Ser Asp Ser Ser Gln
145                 150                 155                 160

gcc atc atc gac aag gtt gcc gcc cag aaa gcc aac atg att gcc gcg      528
Ala Ile Ile Asp Lys Val Ala Ala Gln Lys Ala Asn Met Ile Ala Ala
                165                 170                 175

ata gac cgg aac ctg tcg gag ctg cgg ctc agc cgc cat ggt ttc gtg      576
Ile Asp Arg Asn Leu Ser Glu Leu Arg Leu Ser Arg His Gly Phe Val
            180                 185                 190

ttc gtg gtt gcg gac gat ggc acg gtg atc gtg ccg cca ccc cca tcg      624
Phe Val Val Ala Asp Asp Gly Thr Val Ile Val Pro Pro Pro Pro Ser
        195                 200                 205

gcc gcc cgg ctg ctg gac tcg aca gac gtc gaa tcg gga cgg gta ttg      672
Ala Ala Arg Leu Leu Asp Ser Thr Asp Val Glu Ser Gly Arg Val Leu
    210                 215                 220

cat tcg atg ctt gcc gaa atc tcg tct acc cgc ggc ctg acg ttg cgc      720
His Ser Met Leu Ala Glu Ile Ser Ser Thr Arg Gly Leu Thr Leu Arg
225                 230                 235                 240

ttt acc aac ggc gaa agc gcc tgg cag atc gac gcc ctg cga tac aag      768
Phe Thr Asn Gly Glu Ser Ala Trp Gln Ile Asp Ala Leu Arg Tyr Lys
                245                 250                 255

ccg ctg cat tgg acc atc atc ggt gtc gtt ccc gag ccg gac ctg acc      816
Pro Leu His Trp Thr Ile Ile Gly Val Val Pro Glu Pro Asp Leu Thr
            260                 265                 270

gac ccg gca cag aat ctg gtg cgc cgg cag gca ctg atc ttc gcc gcc      864
Asp Pro Ala Gln Asn Leu Val Arg Arg Gln Ala Leu Ile Phe Ala Ala
        275                 280                 285

acc ttg ctg gcc ggg ctg atg ctg gca tgg gtg gtg gcg gtg cgc atc      912
Thr Leu Leu Ala Gly Leu Met Leu Ala Trp Val Val Ala Val Arg Ile
    290                 295                 300

gcc cgg ccg ttg gcg caa ctg agc aac tac gct cgc cag ctt ccc acc      960
Ala Arg Pro Leu Ala Gln Leu Ser Asn Tyr Ala Arg Gln Leu Pro Thr
305                 310                 315                 320

cag gac ctc acc gag ccg atc cgg gtt ccg ccg tcg gtg gca tgc ctg     1008
Gln Asp Leu Thr Glu Pro Ile Arg Val Pro Pro Ser Val Ala Cys Leu
                325                 330                 335

ccg cgc cgg cgg cgc gac gaa gtc gga cag ctc gcc gaa tcg ttc ctg     1056
Pro Arg Arg Arg Arg Asp Glu Val Gly Gln Leu Ala Glu Ser Phe Leu
            340                 345                 350

ttc atg aac gaa cag ctg cac cac aat gtg cgg gcc ctg atg gcg cag     1104
Phe Met Asn Glu Gln Leu His His Asn Val Arg Ala Leu Met ala Gln
        355                 360                 365

ata tcg aac cgc gaa cgc ctc gaa agc gaa ttg agc atc gcc cgc tcc     1152
Ile Ser Asn Arg Glu Arg Leu Glu Ser Glu Leu Ser Ile Ala Arg Ser
```

```
                370                 375                 380
atc caa ctt ggc ctg ctt ccc cag ccg ttg ccc gat gcg gcc acg cgc      1200
Ile Gln Leu Gly Leu Leu Pro Gln Pro Leu Pro Asp Ala Ala Thr Arg
385                 390                 395                 400

ggc agc cag ttg cgt gcc gtc atg tac ccg gcc cgg gag gtc ggt ggg      1248
Gly Ser Gln Leu Arg Ala Val Met Tyr Pro Ala Arg Glu Val Gly Gly
                405                 410                 415

gat ttc tac gac tac ttc gtg ctg gca gac ggg cgt ctg tgc ttt gcc      1296
Asp Phe Tyr Asp Tyr Phe Val Leu Ala Asp Gly Arg Leu Cys Phe Ala
            420                 425                 430

atc ggc gac gta tcc gga aaa ggc gtg ccc gcg gcc ctg ttc atg gcc      1344
Ile Gly Asp Val Ser Gly Lys Gly Val Pro Ala Ala Leu Phe Met ala
        435                 440                 445

atc gtc agg acc ttg ata cgc agc gtg gcg gaa gaa gag cac gac ccg      1392
Ile Val Arg Thr Leu Ile Arg Ser Val Ala Glu Glu Glu His Asp Pro
    450                 455                 460

ggc gcc atc gcc acc aag gtg aac cac cgt ctg gcc gag aac aac ccc      1440
Gly Ala Ile Ala Thr Lys Val Asn His Arg Leu Ala Glu Asn Asn Pro
465                 470                 475                 480

aag ctg atg ttt gtc acc ttg ctg ata ggc gtc ttc acc ccg gaa aca      1488
Lys Leu Met Phe Val Thr Leu Leu Ile Gly Val Phe Thr Pro Glu Thr
                485                 490                 495

ggc gcc ctg gcc tgg gtc aac gcc ggc cac ccg ccg ccg ctg ctc atc      1536
Gly Ala Leu Ala Trp Val Asn Ala Gly His Pro Pro Pro Leu Leu Ile
            500                 505                 510

gac gaa cgt ggc gag gtc cgc ctg ctt caa gga agc agc ggc gcg gcc      1584
Asp Glu Arg Gly Glu Val Arg Leu Leu Gln Gly Ser Ser Gly Ala Ala
        515                 520                 525

tgc ggc gtg ctg gac aac gag gcg tat tcc acc ctg agc acc acc ttg      1632
Cys Gly Val Leu Asp Asn Glu Ala Tyr Ser Thr Leu Ser Thr Thr Leu
    530                 535                 540

ccg aac ggc acc tcg ctg gtc gcg ttt acc gac ggc gtc acc gaa gcc      1680
Pro Asn Gly Thr Ser Leu Val Ala Phe Thr Asp Gly Val Thr Glu Ala
545                 550                 555                 560

atc cac ggc ggc tgc gcc cag tat ggt ctg ccg cgg ctg gtc gcc ctg      1728
Ile His Gly Gly Cys Ala Gln Tyr Gly Leu Pro Arg Leu Val Ala Leu
                565                 570                 575

atg cag ggc gcg ccg cac gca gcg gcc gaa ctc atc gag cac att ctg      1776
Met Gln Gly Ala Pro His Ala Ala Ala Glu Leu Ile Glu His Ile Leu
            580                 585                 590

cac gac cta cgc gaa ttc gcc gcc gat tcc gaa caa tcc gac gat ctc      1824
His Asp Leu Arg Glu Phe Ala Ala Asp Ser Glu Gln Ser Asp Asp Leu
        595                 600                 605

acc atc atc gcc att cat cgc cca tga                                 1851
Thr Ile Ile Ala Ile His Arg Pro  *
```

610 615

<210> 82
<211> 616
<212> PRT
<213> Bordetella pertussis

<400> 82

Met Asp Leu Val Val Arg Asp Thr Asp Thr Arg Trp Ser Thr Leu Leu
1 5 10 15
Asp Asp Lys Ile Arg Thr Ile Arg Glu Ser Arg Arg Gln Leu Ile Gln
20 25 30
Leu Ser Ala Val Val Thr Ser Val Leu Asn Ala Tyr Ala Ala Gln Ala
35 40 45
Glu Arg Gly His Val Thr Thr Gly Ala Ala Lys Gly Met *ala* Arg Val
50 55 60
Trp Leu Asn His Leu Asp Leu Gly Pro Arg Arg Val Ala Phe Ala Tyr
65 70 75 80
Asp Ala Glu Gly Thr Val Leu Ala Ser Thr Asn Pro Arg Met Ile Asp
85 90 95
Arg Asp Leu Ser Gly Ile Arg Asp Phe Lys Gly Arg Pro Leu Ala Ala
100 105 110
Ala Met Tyr Glu Glu Ser Arg Asn Asp Gly Arg Gly Phe Ala Ile Tyr
115 120 125
Pro Ser Pro Leu Asp Glu Ser Ala Gln Met Arg His Ala Tyr Phe Val
130 135 140
Tyr Phe Pro Ala Trp Lys Trp Val Leu Ala Ile Ser Asp Ser Ser Gln
145 150 155 160
Ala Ile Ile Asp Lys Val Ala Ala Gln Lys Ala Asn Met Ile Ala Ala
165 170 175
Ile Asp Arg Asn Leu Ser Glu Leu Arg Leu Ser Arg His Gly Phe Val
180 185 190
Phe Val Val Ala Asp Asp Gly Thr Val Ile Val Pro Pro Pro Pro Ser
195 200 205
Ala Ala Arg Leu Leu Asp Ser Thr Asp Val Glu Ser Gly Arg Val Leu
210 215 220
His Ser Met Leu Ala Glu Ile Ser Ser Thr Arg Gly Leu Thr Leu Arg
225 230 235 240
Phe Thr Asn Gly Glu Ser Ala Trp Gln Ile Asp Ala Leu Arg Tyr Lys
245 250 255
Pro Leu His Trp Thr Ile Ile Gly Val Val Pro Glu Pro Asp Leu Thr
260 265 270
Asp Pro Ala Gln Asn Leu Val Arg Arg Gln Ala Leu Ile Phe Ala Ala
275 280 285
Thr Leu Leu Ala Gly Leu Met Leu Ala Trp Val Val Ala Val Arg Ile
290 295 300
Ala Arg Pro Leu Ala Gln Leu Ser Asn Tyr Ala Arg Gln Leu Pro Thr
305 310 315 320
Gln Asp Leu Thr Glu Pro Ile Arg Val Pro Pro Ser Val Ala Cys Leu
325 330 335
Pro Arg Arg Arg Arg Asp Glu Val Gly Gln Leu Ala Glu Ser Phe Leu
340 345 350
Phe Met Asn Glu Gln Leu His His Asn Val Arg Ala Leu Met *ala* Gln
355 360 365
Ile Ser Asn Arg Glu Arg Leu Glu Ser Glu Leu Ser Ile Ala Arg Ser
370 375 380
Ile Gln Leu Gly Leu Leu Pro Gln Pro Leu Pro Asp Ala Ala Thr Arg
385 390 395 400
Gly Ser Gln Leu Arg Ala Val Met Tyr Pro Ala Arg Glu Val Gly Gly

405 410 415
Asp Phe Tyr Asp Tyr Phe Val Leu Ala Asp Gly Arg Leu Cys Phe Ala
420 425 430
Ile Gly Asp Val Ser Gly Lys Gly Val Pro Ala Ala Leu Phe Met ala
435 440 445
Ile Val Arg Thr Leu Ile Arg Ser Val Ala Glu Glu Glu His Asp Pro
450 455 460
Gly Ala Ile Ala Thr Lys Val Asn His Arg Leu Ala Glu Asn Asn Pro
465 470 475 480
Lys Leu Met Phe Val Thr Leu Leu Ile Gly Val Phe Thr Pro Glu Thr
485 490 495
Gly Ala Leu Ala Trp Val Asn Ala Gly His Pro Pro Pro Leu Leu Ile
500 505 510
Asp Glu Arg Gly Glu Val Arg Leu Leu Gln Gly Ser Ser Gly Ala Ala
515 520 525
Cys Gly Val Leu Asp Asn Glu Ala Tyr Ser Thr Leu Ser Thr Thr Leu
530 535 540
Pro Asn Gly Thr Ser Leu Val Ala Phe Thr Asp Gly Val Thr Glu Ala
545 550 555 560
Ile His Gly Gly Cys Ala Gln Tyr Gly Leu Pro Arg Leu Val Ala Leu
565 570 575
Met Gln Gly Ala Pro His Ala Ala Ala Glu Leu Ile Glu His Ile Leu
580 585 590
His Asp Leu Arg Glu Phe Ala Ala Asp Ser Glu Gln Ser Asp Asp Leu
595 600 605
Thr Ile Ile Ala Ile His Arg Pro
610 615

Amended Claims for PCT Application No. PCT/EP99/10297

Claims:

1. An isolated polypeptide comprising an amino acid sequence which has at least 75% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 and 72 over its entire length.

2. The polypeptide as claimed in claim 1 comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 and 72.

3. An isolated polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72.

4. An isolated polypeptide comprising a fragment of at least 7 consecutive amino acids of the polypeptide as claimed in any one of claims 1 to 3, wherein the fragment comprises an epitope.

5. The polypeptide of claim 4, wherein the fragment is immunogenic.

6. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 75% identity to the amino acid sequence of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72 over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

7. An isolated polynucleotide comprising a nucleotide sequence that has at least 75% identity to a nucleotide sequence, encoding a polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72, over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

8. An isolated polynucleotide which comprises a nucleotide sequence which has at least 75% identity to that of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length; or a nucleotide sequence complementary to said isolated polynucleotide.

9. The isolated polynucleotide as claimed in any one of claims 6 to 8 in which the identity is at least 95% to SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 over its entire length.

10. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72.

11. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71.

12. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70 or 72, obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69 or 71 or a fragment thereof.

13. An expression vector comprising an isolated polynucleotide according to any one of claims 6 - 12.

14. A recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 6 - 12.

15. A host cell comprising the expression vector of claim 13 or a subcellular fraction or a membrane of said host cell.

16. A process for producing the polypeptide of claim 1 comprising the steps of culturing a host cell of claim 15 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

17. A process for expressing a polynucleotide of any one of claims 6 - 12 comprising transforming a host cell with an expression vector comprising at least one of said polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

18. A vaccine composition comprising an effective amount of the polypeptide of any one of claims 1 to 5 and a pharmaceutically acceptable carrier.

19. The vaccine composition of claim 18, wherein the polypeptide has an amino acid sequence selected from the group consisting of: SEQ ID NO:42, 46, 48, 50, 52, 54, 56, 58, 60 and 62.

20. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 6 to 12 and a pharmaceutically acceptable carrier.

21. The vaccine composition according to any one of claims 18-20, wherein said composition comprises at least one other *Bordetella pertussis* antigen.

22. An antibody immunospecific for the amino acid sequence of claim 1 or 2, the polypeptide of claim 3 or the fragment of claim 4 or 5.

23. A method of diagnosing a *Bordetella pertussis* infection, comprising identifying a polypeptide as claimed in any one of claims 1 - 5, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

24. Use of a composition comprising an immunologically effective amount of a polypeptide as claimed in any one of claims 1 – 5 in the preparation of a medicament for use in generating an immune response in an animal.

25. Use of a composition comprising an immunologically effective amount of a polynucleotide as claimed in any one of claims 6 - 12 in the preparation of a medicament for use in generating an immune response in an animal.

26. A therapeutic composition useful in treating humans with *Bordetella pertussis* disease comprising at least one antibody directed against the polypeptide of claims 1 – 5 and a suitable pharmaceutical carrier.

27. A kit for diagnosing infection with *B. pertussis* bacteria in a human comprising a polynucleotide of claims 6-12 or a polypeptide of claims 1-5.

28. A method of identifying virulence genes from a pathogenicity island containing a type III secretion system from pathogenic strains of bacteria, comprising:

designing degenerate PCR primers complementary to well-conserved regions specific to the LcrD polypeptide of *Yersinia*;

amplifying the polynucleotide containing the DNA sequence between (and including the DNA sequence of) the primers of *lcrD*-like genes present in said pathogenic strain of bacteria;

sequencing the *lcrD*-like gene;

determining whether the DNA sequence is more homologous: to the virulence-associated family of *lcrD*-like genes, or to the flagellar-associated family of *lcrD*-like genes; and

if a virulence-associated member, sequencing the entire pathogenicity island, and

identifying genes within this sequence.

29. A method of determining whether a particular bacterial strain harbours a type III secretion system involved in pathogenicity, comprising:

designing degenerate PCR primers complementary to well-conserved regions specific to the LcrD polypeptide of *Yersinia*;

amplifying the polynucleotide containing the DNA sequence between (and including the DNA sequence of) the primers to determine the presence of any *lcrD*-like genes in said bacterial strain;

if amplified successfully, sequencing the *lcrD*-like gene; and

determining whether the DNA sequence is more homologous: to the virulence-associated family of *lcrD*-like genes, or to the flagellar-associated family of *lcrD*-like genes.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C07K 14/235, C12N 15/31, 5/10, A61K 39/10, C07K 16/12, G01N 33/569, C12Q 1/68

A2

(11) International Publication Number: WO 00/37493

(43) International Publication Date: 29 June 2000 (29.06.00)

(21) International Application Number: PCT/EP99/10297

(22) International Filing Date: 21 December 1999 (21.12.99)

(30) Priority Data:
9828217.1 21 December 1998 (21.12.98) GB

(71) Applicant *(for all designated States except US)*: UNIVERSITE LIBRE DE BRUXELLES [BE/BE]; Faculté des Sciences, Avenue F. Roosevelt, 50, B–1050 Bruxelles (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BOLLEN, Alex [BE/BE]; Université Libre de Bruxelles, Faculté des Sciences, Avenue F. Roosevelt, 50, B–1050 Bruxelles (BE). FAUCONNIER, Alain [BE/BE]; Université Libre de Bruxelles, Faculté des Sciences, Avenue F. Roosevelt, 50, B–1050 Bruxelles (BE). GODFROID, Edmond [BE/BE]; Université Libre de Bruxelles, Faculté des Sciences, Avenue F. Roosevelt, 50, B–1050 Bruxelles (BE).

(74) Agent: TYRRELL, William, Arthur, Russell; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) **Title:** VACCINE

(57) **Abstract**

This invention relates to a general method for detecting pathogenic strains of bacteria which harbour a type III secretion system. More particularly, this invention relates to the methods as applied to the pathogen *Bordetella pertussis*. Furthermore, the invention relates to newly identified polynucleotides within these regions, virulent polypeptides encoded by them and to the use of such polynucleotides and polypeptides, and to their production. More particularly the polynucleotides and polypeptides of the present invention relate to the virulent effector proteins associated with the type III secretion system of *Bordetella pertussis*, which are particularly suitable for vaccine purposes.

Fig. 1

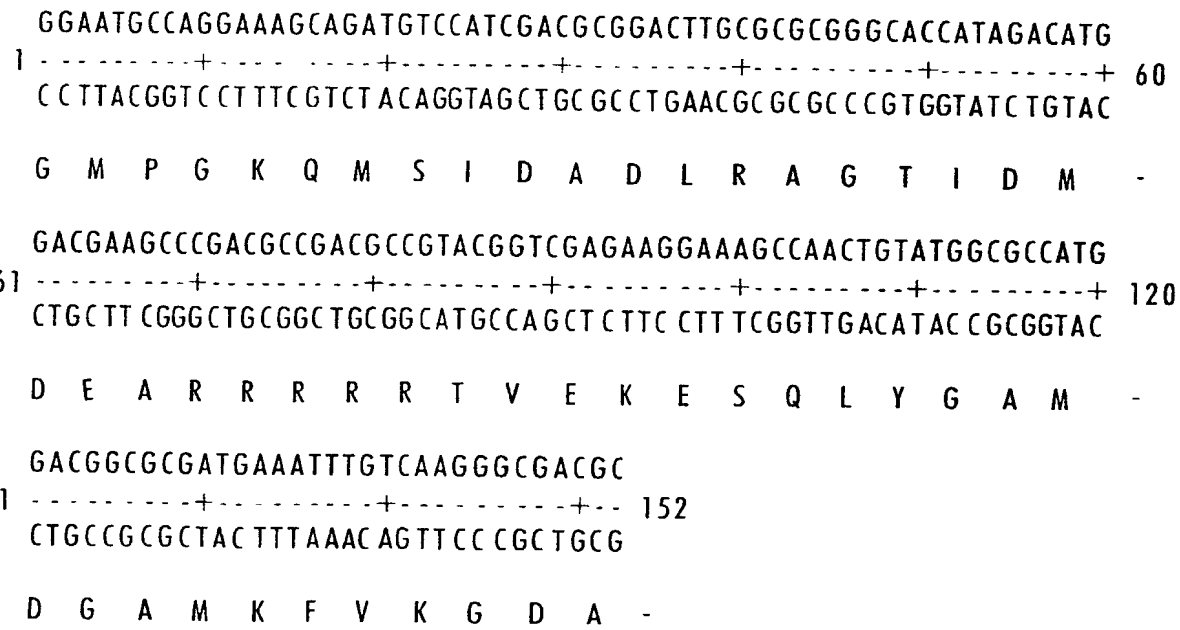
GGAATGCCAGGAAAGCAGATGTCCATCGACGCGGACTTGCGCGCGGGCACCATAGACATG
1 ---------+---- ----+---------+---------+---------+---------+ 60
CCTTACGGTCCTTTCGTCTACAGGTAGCTGCGCCTGAACGCGCGCCCGTGGTATCTGTAC
G M P G K Q M S I D A D L R A G T I D M -
GACGAAGCCCGACGCCGACGCCGTACGGTCGAGAAGGAAAGCCAACTGTATGGCGCCATG
61 ---------+---------+---------+---------+---------+---------+ 120
CTGCTTCGGGCTGCGGCTGCGGCATGCCAGCTCTTCCTTTCGGTTGACATACCGCGGTAC
D E A R R R R R T V E K E S Q L Y G A M -
GACGGCGCGATGAAATTTGTCAAGGGCGACGC
121 ---------+---------+---------+-- 152
CTGCCGCGCTACTTTAAACAGTTCCCGCTGCG
D G A M K F V K G D A -

Fig. 2

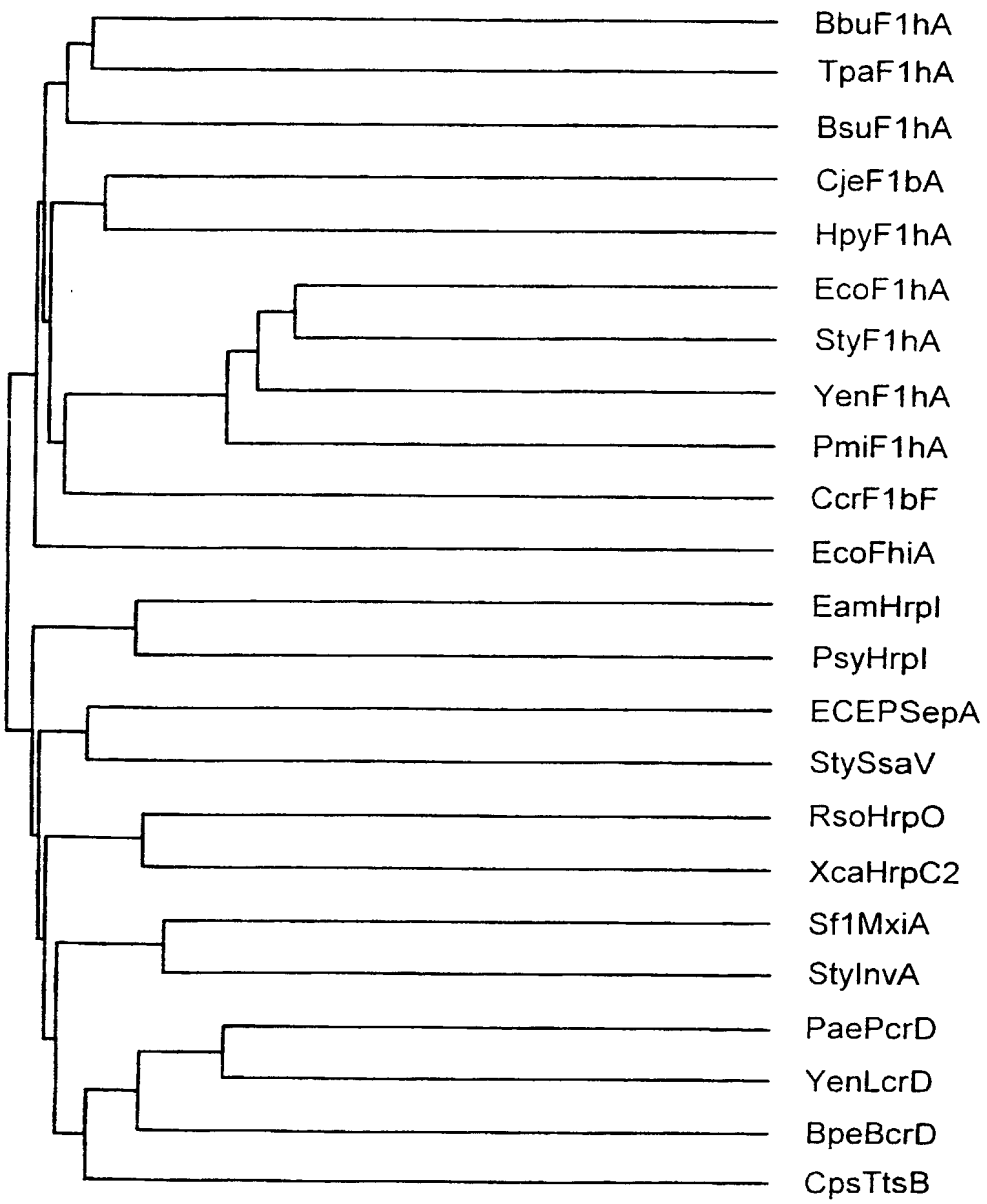
BbuF1hA
TpaF1hA
BsuF1hA
CjeF1bA
HpyF1hA
EcoF1hA
StyF1hA
YenF1hA
PmiF1hA
CcrF1bF
EcoFhiA
EamHrpI
PsyHrpI
ECEPSepA
StySsaV
RsoHrpO
XcaHrpC2
Sf1MxiA
StyInvA
PaePcrD
YenLcrD
BpeBcrD
CpsTtsB

Fig. 3

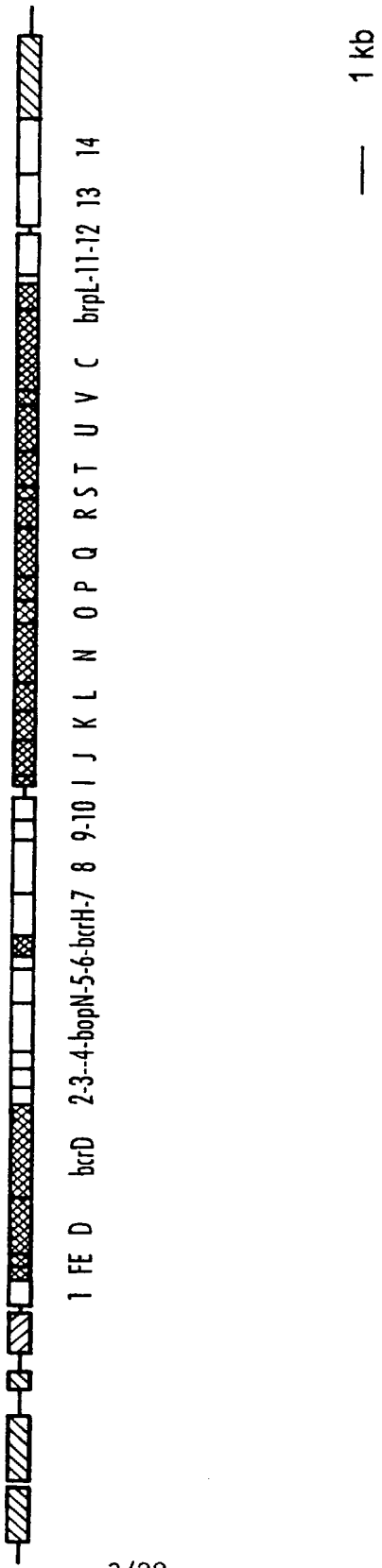
1 FE D bcrD 2-3--4-bopN-5-6-bcrH-7 8 9-10 I J K L N O P Q R S T U V C brpL-11-12 13 14
1 kb

Fig. 4

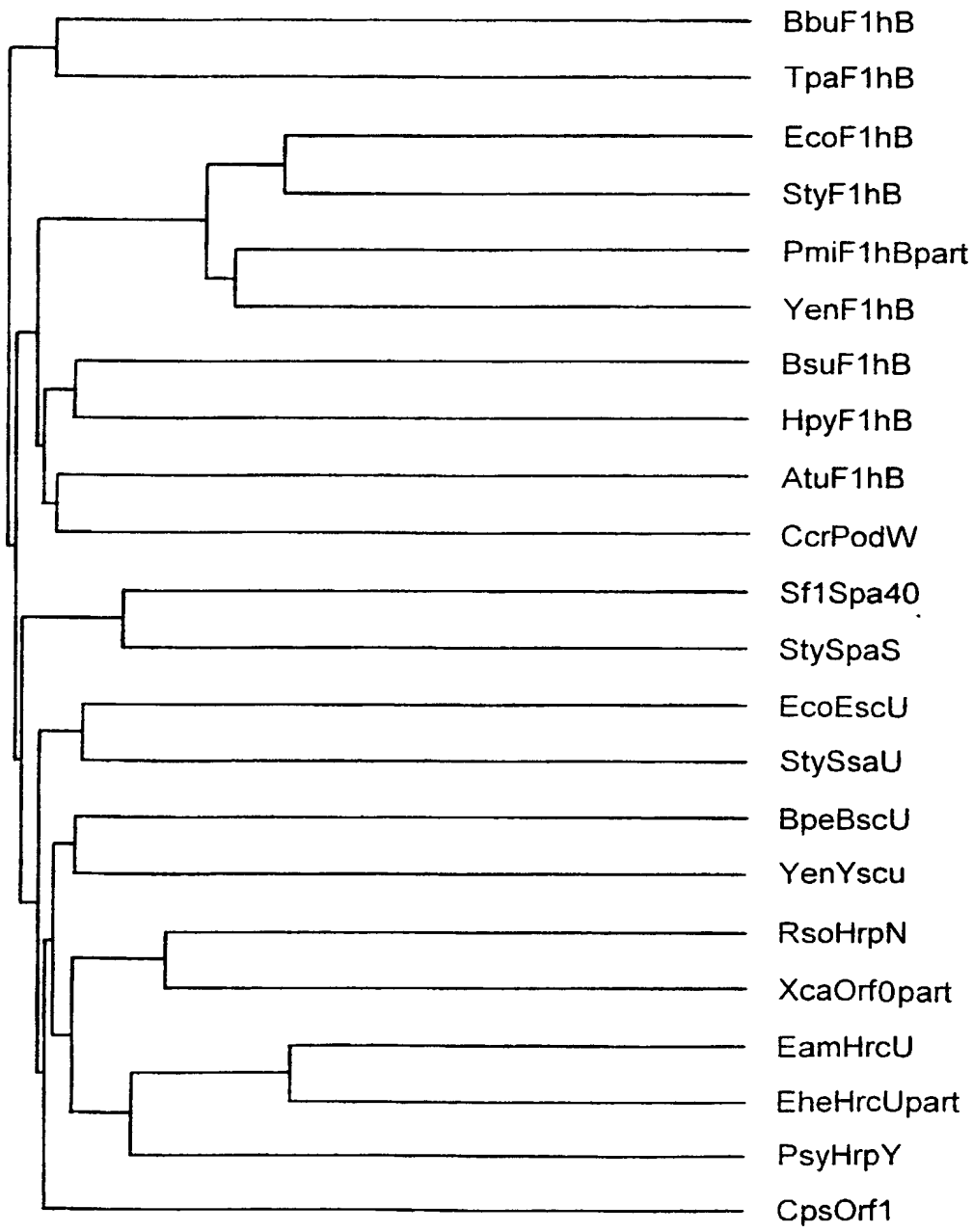
BbuF1hB
TpaF1hB
EcoF1hB
StyF1hB
PmiF1hBpart
YenF1hB
BsuF1hB
HpyF1hB
AtuF1hB
CcrPodW
Sf1Spa40
StySpaS
EcoEscU
StySsaU
BpeBscU
YenYscu
RsoHrpN
XcaOrf0part
EamHrcU
EheHrcUpart
PsyHrpY
CpsOrf1

Figure 5

```
   1  GATCTCCAGC TTGATGTCCG GATGGGCCTT CTCGAAGGCC GCCTTGTAGG
  51  CCTGGGTCAG GTCCTTGGGG AAAGACGTGA TGACCGTTAC CGTGCCGGCC
 101  AGGGCCGGCG CGCAGGCAGC CATGATGGCC GCTGCGAGGA CTGGGCGCAA
 151  CGCTTGCATG GGTCTCCTCC TTTCTTGAGT TGTGGCAAGA CCTTAATGCC
 201  CGTGTTCCCC CGAGTCCAAT CAGAAATTTT GATGCGCGCC ATCAACGGCG
 251  CGGCCGGCTT CAGCGCGCCG CCAGCCGCTG CGCCAGGCGC AATTGCTCTT
 301  CCAGCAGCTT CTGGATCTGC CCCTGGATGG CGAAGGCCTG GCGGTTCAGG
 351  TCCTGCAGCT GCCGGGCCTG CTGCTCGGCG GTGGCGTTGC TTTCGCGCAC
 401  CTGGCGCATC TGGCGCATGA TGGCCGCCAG CTGGCGCTGC AGGGCCTGCA
 451  AGGCCTCGGC CGCGGGGCTG CCGGCCTCGT CCTTTGCGCC GCTGGCGGCG
 501  GCCCGGTCGC GCAACTGCCT GGCGTACTGG TCGAACGCGC CGGCGGAGAT
 551  GCGCGATGAA ACGGAACTGA CGGTCATGGT GGCGGCGTCC GCGTGGATTA
 601  GGGTTAGTCC TTAATACGGC AGGCGCGCGG ACGAACTTAA ACATATTGCT
 651  GCGCGCTTCC GACCCCCCTT ACACTTGCGC CGCCCATCCC TTCAACTTCC
 701  ACGCATACCT ATGTCCAATA CCTATTTCCC GCGCTGGCGG CTGGCCGACG
 751  ACACCGTGCC GGGCGCGGTC ATCGCGCCCG ACGAACGCCT GTCCTGGCCC
 801  AAGAACATCG CCATGGGGGC CCAGCACGTG GTCGCCATGT TCGGTTCCAC
 851  CGTGCTGGCG CCGCTGCTGA TGGGTTTCGA CCCCAATGTG GCGATCCTCA
 901  TGTCCGGCAT CGGCACGCTG ATCTTCTTCC TGTTCGTCGG CGGCCGGGTG
 951  CCCAGCTACC TGGGCTCCAG CTTCGCCTTC ATCGGCGGGG TGGTGGCGGT
1001  CACCGGCTAT GTGGCGCCCG GCGCCAACGC CAATATCGGC GTGGCGCTCG
1051  GCGCGATCAT CGCCTGTGGC CTGGTGTACG CGCTGATCGG CCTGGTCGTA
1101  TGGGCGGCCA GCGCGCGCGG CAACGGGGCG CGCTGGATCG AGGCCATGAT
1151  GCCGCCGGTC GTCACGGGCG CGGTGGTGGC GGTGATCGGC CTGAACCTGG
1201  CCCCGATCGC CGCCAAGGGC GCCATGGGTT CGTCCGGCTT CGAGGCCAGC
1251  ATGGCGTTGA TGACCATCCT GTGCGTGGGC GGCATCGCCG TCTACACGCG
1301  CGGCATGGTG CAGCGGCTGC TGATCCTGGT CGGCCTGGTG CTGGCCTGCG
1351  TCATCTACGC GGTCTGCGCC AACGGCCTGG GGCTGGGCGC GCCCATGGAC
1401  TTCGCCAAGG TGGCCGCCGC GCCGTGGTTC GGCCTGCCCA GCTTCGCCGC
1451  GCCGGTGTTC GAGCCGCAGG CCATGGGCCT GATCGTGCCG GTGGCCATCA
```

Figure 5 (continued)

1501 TCCTGGTGGC CGAGAACCTG GGCCACGTGA AGGCGGTCGC CGCCATGACC

1551 GGACAGGACC TGGACCGCTA CGTGGGCCGC GCCTTCGTGG GCGACGGCGT

1601 GGCGACCATG GTTTCCGGCG CCGTCGGCGG CACCGGGGTG ACCACCTACG

1651 CCGAGAATAT CGGCGTGATG GCCGTGACGC GCATCTATTC CACGCTGGTG

1701 TTCGTGGTGG CGGCCGTGAT CGCGCTGGTG CTGGGGTTCT CGCCCAAGTT

1751 CGGCGCGCTG ATCCAGACCA TCCCCGGCCC CGTGCTGGGG GGCATGTCGG

1801 TCGTGGTGTT CGGCCTGATC GCCATCGCCG GCGCGCGCAT CTGGGTGGTC

1851 AACCAGGTCG ATTTCAGCGA CAACCGCAAT CTGATCGTGG CCGCCGTGAC

1901 CCTGGTGCTG GGGGCGGGCG ACTTCAGCGT CAAGCTGGGC GATTTCTCGA

1951 TGAACGGCAT CGGCACCGCC ACGTTCGGCG CCATCATCCT GTACGCCCTG

2001 CTGGGCCTGG CGCGTCGCCG CTGACGGCGC GCCAACCCGG CACCGGGGCC

2051 GCGCTCAGTG GGCCGCGGCC GGCGGCAGGG CGGCATGGCC GCCCGCCGGC

2101 CCGGCCAGCG GCGCGGCGGC GACGTCGATG ATGGCGCCGC CCTGCGTCCT

2151 GAAGACCGCC GCCGCGGCCG CCAGGTGGCG GGCCTGTTCT TCCATGGCCG

2201 TGGCCGCCGC CGCGGCCTGT TCGACCAGCG CGGCATTCTG CTGGGTGGTT

2251 TCGTCCATTT GCGAAATCGC CAGGCTGACC TGGTCGATGC CGCTGGCCTG

2301 CTGGGCCGAG GCGGCCGAGA TCTCCCCCAT GATGTCGGCC ACGCGCTGTA

2351 CCGAGGCCAC CACCTCGTCC ATGGTGCCGC CGGCGCTGGC GACCTGTTGC

2401 GAGCCGGCGC GCACCGTCGC CACCGAGCTC TCGATCAGGG CCTTGATCTC

2451 CTTGGCCGCC TGCGCGGCGC GCTGGGCCAG GCTGCGCACC TCGCCCGCCA

2501 CCACCGCGAA GCCCTTGCCC TGTTCGCCGG CGCGCGCCGC CTCGACCGCG

2551 GCGTTCAGCG CCAGGATATT GGTCTGGAAG GCGATGCCGT CGATCACGGT

2601 GACGATGTCG GCGATCTGGC GCGAGCTCGC GGAGATGCCG TGCATGGTCT

2651 GCACCACCTG CGCGACCGAC TCGCCGCCGC GCTGCGCCAC CTGCATGCTG

2701 ACGGCGGCCA GCTGATTGGC CTGCGCGGCA TTGTCGGCGT TCTGCTTGAC

2751 CGTGGTGGCC AGTTCCTCCA TGGTGGCCGC GGTCTCTTCC AGGGCGGCGG

2801 CCTGCTCCTC CGTGCGGCTG GACAGGTTGG CGTTGCCGGC CGAGATCTCG

2851 GCCGCGCCGA CGTTGATTTC GTCGACGCCG CGCCGCATGA CGGCGATGGT

2901 GCGCGTCAGG CCTTCCTGCA TGCGCTTGAG CGCCGCGAAC AGCGCGCCGA

2951 TTTCATTGGC CGAGCGCACC TCGATGCGCG CGGTGAGGTC GCCGTCGGCG

3001 ATGCGGTCGA AATGATGGCC GGCCTCCAGC AAGGGGCGCA GCACCGCGCG

Figure 5 (continued)

```
3051  GCGCACGAAC AGCCAGCCGG CCAGGGTCAG CAGCACGCCC AGCGTGGTCA
3101  GCGCGATGGC GCTCCAGCGG GCCACGACAT GGGTGTCCTC GGCGCCGCTG
3151  CGCACTTCGT CGCTGTGCGC CTGTACCCTG GCCAGAAAGG CTTCCATGTC
3201  GCGCTGGAAT GCATGTTCGG CCTGCTGCGC GCGATGCATG GCGGCCAGGG
3251  CGGGCTCGGC CTGGCCGGCG TCGACGGCGG CGGCCAGCTC GTCCAGCACC
3301  GACTGGTAGG CGCGCCAGCG CGTCTGCAGG GTCGCGGCCA GCTCGGCCGC
3351  GGCCGGCTCC TTGGGCACGT CCACGTAGCG CTGGAATTGC GTGGCCGCCT
3401  GGCGCTGCGC CGCGCCGCTG GCCTCGAACA GCGGGTCGAC CTGGTTGACG
3451  GCGACCTGGT TCAGCCCCTC GATCTCGGCG GCGCTCCTGC CCGCGTTGCG
3501  CCACGTCAGG GCGCCCGACA GCATCAGCAC CGCCAGGAAG CACGCGAATA
3551  CCGCCACCAT CGCGGTGCGG ACCTTGATCC TTTCAAGCAT GGGCATCTCC
3601  TGGGTAGGGT CTTGCGCAAT AAACGTCGAC GACGTGCGCG GCGTACCCGC
3651  GCGGGAAGTG TCAGGCCGTG ACGTGGAATC CGTTGTTGAT GTGGAATTCG
3701  ACTTAAATGA TCGGTATTCG GTATTAATTT GTTATTTGTA GTTATATATA
3751  CGAATATTCA TACCCGGATT TGCCCTAAGT TGGTGCGTTC TCGACGGGTG
3801  CTTTCGATTG CCCGGGGCCT GCGGCGCATG AAGGAATGCG CGGCAACGCC
3851  GGGCCCGCGC TGCGATCGGC GGCCAACACG CAGGTTTTGT GGCTTTTCCG
3901  CAGCCAACAT GCGCCGAAAC CTACGCCGGG CTTACAGGCT TGCAATTCCG
3951  GTGGACTTTG CCGACAATGT CATCTGATTG CCCCGGTTCC GACCCGAGCC
4001  GGGGTTTTGT TTTGGTCGAC GCTTGGCCGC CGGATGCGGC AGGCCGATCA
4051  AAGAGGAGAC AGCAAAGGGG AGCCTCGGTC GGGTTCGACC TTGTCTCGTC
4101  TTTTGTTCGC CTGTCTTCCG CAGCGGCCCG CCTGTTTCAT GGCGAGGCCA
4151  TACACCAAGC CGAGACCTTC ACCGAAACGC TCCGTCGGGG GCGTTTTCTA
4201  CTTTTGTTTG GGAAACGACA TGTCTGCCAT TCCTTTGACC GTGCGCGGGG
4251  CCGAGCGCTT GCAGCAAGAA CTGCATCGGC TTAAGACCGT TGAGCGTCCT
4301  GCGGTGATCA GCGCCATTGC GGAGGCGCGT GCGCAGGGTG ATTTGTCGGA
4351  AAATGCCGAG TACGACGCCG CCCGCGAACG CCAGGGCTTC ATCGAAGGCC
4401  GGATCTCCGA ACTCGAGGGC ACGCTTTCGA ACGCGCACCT CATCGATCCA
4451  ACGGCGCTCG ACGCCGAAGG CCGTGCCGTG TTCGGCGCGA CCGTGGAAAT
4501  CGAAGACCTC GACTCGGGCG ACCGCCTGAC CTACCAGATC GTGGGCGACG
```

Figure 5 (continued)

4551 TCGAAGCCGA CATCAAGTCC AACCTGATTT CGGTCTCCAG CCCGGTGGCC

4601 CGCGCCCTGA TCGGCAAATC CGAGGGCGAT GTGGTCGAAG TGAAGGTGCC

4651 GGCTGGCGTG CGCGAGTACG AAGTCATCGG TGTGCGTTAT CTCTGACGCC

4701 GATTCCGCCC CCCTGCATAC CCATGGCCAA TGACCGACGC CGTTTCCACC

4751 ATCAGCCAGC CGCTTGCGGC CAGGCGGGGT GGCACAGCCG CAGCCGTCGC

4801 GCTGCAGGCG TCCCGGATGC GATATCCCGG GCTCCGGTCG TGGCCAGCGT

4851 ATTCCCCTTC TCGGCCGTCA ACGGCTTGGG GGGCGCGGTA TTCATCCGAA

4901 TTCCTGAACG CCGGGCCGTG CGCCGGGCGT CTGTATCGTT GTCGCGCCGC

4951 GATGCACGGA ACGTGCCGTC GGGCGCCGCA ACGCCGGGCC GCGCCGCCTA

5001 GGTGTGAACT GTCAATAGGT TGTATTCGTC CAGGTTGAGT CTGGAGATGG

5051 GTACAGCGCG CCCGATGCCT TGGTGGGGTC GATGCCAGTT GTAGTGGTGT

5101 AGCCAGGATT TCATGGCATC GGCTCGGTGT TGGGAGTTCT GGTAGGTGTG

5151 AGCGTAAGCC CACTCACGCA AGGCCGACTG GATGAAGCGT TCGGCCTTGC

5201 CATTGGTCTG TGGGCGGTAA GGTCGGGTAA AGCGGTGCTT GATGCCCAGC

5251 TCATGGCACA GCGCGGCGAA GGCGCGGCTG CGAAAGGCCG AGCCATTGTC

5301 GGTGAGCAAG CGCTGGATGG TCACGCCCAG GCGCTGGTAG TAGGCCACTG

5351 CGTCCTTGAG GAACTGGACG GCGCTGGGGA AGCGCTCGTC GGGGTGGATG

5401 TCGGTGAAGG CCACGCGGGC GTGGTCATCG ATGGCCACGA AGACGAAGTC

5451 CCAGCCGGCC CCCTCAACGG TATCGCGTCG GTTGCCCGTG ACCCGGTGGC

5501 CAGGGCGCTG GATACGTCCC AGCTTCTTGA TGTCGATGTG CAGCAGATCG

5551 CCGGGGGCCT GATGCTCGTA GCGCACCACC GGCTCGGCCG GCTCCAGGTC

5601 GGCCAGGTGC GACAGACCGG CGCGGGCCAG GACGCGGCTG ACGGTGCTGG

5651 CTGACACGCC CAGCGCCTGG GCGATGCGCG CTTGGGTCAG CCGCTTGCGG

5701 CGCAGCTCCA CGATAGCCAG CGCCTTGGCC GGCGCAATCG CTCGGGGCGA

5751 GACCGTCGGG CGCGAGGACG CATCGGCCAA GCCCGCCTGG CCCTGAGCCA

5801 GGAAGCGGCC CAGCCATTTG CGCACAGTCG GCGCGGTGAC CCCATAGGCG

5851 CGGGCCGCTT CAGGCACACA AACTTGATGG GCGATCAATT GCTGGACCAT

5901 TTCGAGTCGA CGTAGGAAGG TCAATCGGGC ATGCTTATGG GTGTTCATCC

5951 GGCCGGGCTC CTTGAGTGAA CTGGGGGGGT GGCGATTTCC AGTTTCTCAA

6001 ATCCGGTTCG GATGAACCAT GCATACAACC TATTGAATCT TCACAACTAG

6051 CGCGCGTGGC GCGGAAAGAC CAGCAGGTCG GCCGTCACCG GTTCCCTGTT

Figure 5 (continued)

6101 GTCGAATAAC AGGTAATGCG TGGCGTCCGG ATATTCGCCA TGGCGCGCGC

6151 CGGATCCGCA TGCCCTTGGC GGGGTCCTGC GGCCGGACGT CCGAAAACCG

6201 GGATATTCCG AGAAGCAATC GGCTGGCGCC TTCCAGACTG TGCGCATACC

6251 ATTGCCCTCT TTTGCCACGC ATTTCGAGCG TATGGTTTCC CTTGCGCCCG

6301 ACGCAGGCTC GCCTCGCCAT GACCGACACG GCATACCACC AACTCATCGC

6351 CGATTTCGGC CGCCTCATCG GCATCGACTC GCTCAACCCC GGTGCCGGCG

6401 GCCTGTGTCA GTTGATTTTC GAACCGTGCG CACCGGTCTT CATCGCACCG

6451 GTGCACGCCC GGACGGAAAT CATGATTTCC TGCGTGCTGG GCACGGCGGA

6501 CGCGGCCAAC CCGGCAAGCA TGGCCCGAGC CAACTTCATG CAGGCCGGCA

6551 GCGGCGTCGT GGCCTGCATC GGCGGCGATG GGTTGTTCTA TCTGCAGCAG

6601 GCCATACCCC TGTCGCGCGC CACGCCCGCA ATCCTGCTCG ATCACTGTGA

6651 GCGTCTGCTG CAGGAAGCCT CGCGCTGGCG CGTCGGCGAC CACGACGGCT

6701 GCGCCACCTC GGCCCCGAAT ATCGCCGCGC TGACGCGCGG CGTCTAGGTC

6751 GGCGGCGACT GCCGGGCGCC GCCGCGGCAG GCTCAACTCG CCTTCTGTAT

6801 GACGCCCTTG AGCGAATCCG CGACCTGCTT GACCACCGTG CTCTGTAGAT

6851 CGATCATCAC GACCCACGAT TGCATTTCCT GTTGCACGAT CAGCAGGTCG

6901 GACGTGCTGA CCGCGCCGTC TCCGCGCGCG CTGAGCGCCT CCAGGCGGCT

6951 GCGCAGGTCG CGTTCGTGAG CGTTCAGCCG CGTATTGACC GCCTGGTTGA

7001 CGCTCTGCAT GGTCACTCGG CCTGCGTCGC CTCCCAGGTT AATGGCCATG

7051 CTTGTCTCCT TCGGCGCATT GTTCATTGCT CAGGCGCGTC AAGACTGACG

7101 CCGGAGGGTT GTCCGGCCCG GTCGGGCGCT GCAGCAATAC CTGCCGGGCC

7151 GCGGTGGCGG CCGCCAGCGC GTCGCGCCAT ACGGGATAGG TGTCGCGCCC

7201 CACGCCATCG TGCAGGCGGA CGCGCAAACG TGTCTCCAGT TCGCCAAGCT

7251 GCGACAGCAA GGTGTCGCGC AAGGCGGAAC CGCCCGGCGA TGCCAGGCGC

7301 ACTTCCAGTT CGGTCAGGGC TAGAACAGAT GTACTCATAT GATGTTGCAG

7351 ACGAGGGTTG ACGGCTACCG AGGCGATTTC ATCGCGTCAC GATGACCGGT

7401 TCGGGACCAT CGAAGACCAG GCGGCCGGAT TCGATGGCGG TAAGGCGATA

7451 CTGGTCCCGC AGTCCGCCCA CCAGGAGGCG GCTGCCATCG GCCAGCATCA

7501 GGTACGGTTG CGGGCCGCTC ACGACACTGC GTATCTCGAA CGGCACGTGA

7551 TCCCGCGTCG CGCGCGCGGC GGTGGCCGGC AGCCGGACGA CGTCGTAGTT

Figure 5 (continued)

7601 GCGCTGGTTG AACGCGGCCA CCAGCTCGCG CAGGCGCGCC ATGCGGCCTG

7651 CCGCCAATCC GCCGGGATCT GCGTCCAGGC GGTCGGCGTG CCAGCTGAGC

7701 TTGACGCCGT CGAGGCGTTC GTCGGCCAGC TGGGCCGCGA ACTGGGCCGA

7751 GACCTCGTCG GCCAGGCGTA CATCGCGACC GAGGATCGTC ATGCCCGGCA

7801 GGCGCATGCG CACCGCATGC AGCGCCGCGG CGCGTTCGTG CGCATCGCTG

7851 GCGATGCCCG AGATCGCCAG GCGGCCATTG CCGTACGGGC GCGCCATGTA

7901 GCGCACCCCG AATGTCGCCA GGACATCGCA GGCCAGGGCC CTGGCCTCGT

7951 CCTGCCTGCT TACCTGCATG GCAGGCCGTG GCGCAAGTTG CGCCAACGCC

8001 CTGGCGACCC GAGCGAATTC CGTCTCGTCG TGCACCCATC CGGTCACGGT

8051 GAGCACGCCG CCACGGCCGT AGGCCGCTTG TAATTGCTCG GTAAGGCCCA

8101 GGCTGTCGAT GAGCGCCGCG GCGCGGACCA GCGGCGCGGT GGGCGTTGGG

8151 GGCGGCGCGG CCGGCGGCGT GGCGGGTGTG GTCACGGAAA CCAGCGCCGT

8201 GGCCAGGCCG ACCAGCAGGA CGGCCGCGGC CGCGCCCAGC GCCAGCCAGG

8251 GCCGTCCTGC ACGTCGGCGC GGCATGAGGG CAGCGACGGA CGGCGGGCTT

8301 GTCGAGCCAG GGACGTCGTG CAAGGCTGTG TCGCTGCCGT CCGGGCCGCA

8351 CGGCTCCGGC GGCGCGGGCC ACGGCGCGGA AGGGGCGGCC ACGGTGATCC

8401 AGGCGGCTCC CAGCTCTACG GGTTCGTTGA AGGCCGCGGG CGGACACGGC

8451 GCCTGGGCGT CCAGGCCGGG CGTCACGGCG CCGGCCAACC GCCAGCCGGA

8501 CTGGTCGATC TCCAGCCATC CCGCCACTTC GGGCATGTCC TCGCCGGTCA

8551 GGACGATATC GCAATGCGGA TTGGCGCCCA CGCGCGCGCC ATGCACGGCC

8601 GGGCAGCGCG CCATGCACTG TGCGCCTGAA AGCACGCGGA ATTCCAGCGC

8651 CGTCGTCATA GATCCACCCT GCCCAGGGGC TGTACATTGA TCTCCGGCGT

8701 CAGTTCCTGG TAGGACAGCA CCGGCAGGGC GTAGAGATCG GCTTCTATCA

8751 TCTTGCGCGT GTAGCGCCGG ATGTCCATCG ACGTCAGCAA GACGGGACGG

8801 CTCGCGCCGG CGGCCAGATC GCCGACACAT TGACGGATGT GCTCGACCAG

8851 TCGGCGTGTC GTGTCCGGAT CGAGGGCGAG ATAACTGCCG GCGGCGGTCT

8901 GCCGGATGGC GGCGCGCACG GTTTCCTCGA CCTTGGGGGC CAGCAGGTAG

8951 GCGGGCAGGA TATTGTGGCC GCTGGTGTAC TTGTGGCTGA TATAGCGCTT

9001 GAGTGCGATT CGGACATACT CCGTAAGCAG GACGGTATCC TTTTCCTTCT

9051 GGCCCCATTC GACCAGCGCT TCCAGGACGG CGCGCAGGTT GCGTATCGAC

9101 ACTTCTTCGG AAACAAGGCG CTGCAGGATT TCGGCAATCT TCTGCACCGG

Figure 5 (continued)

```
 9151  CATGACGCGC AGGCACTCCT TGACCAGATC GGGAAATCGT TCTTCCATGG
 9201  CCGAAAGCAG AAACCGGGTT TCCTGGATGC CGATGAAATC GGCTGAATAT
 9251  TTTTTCAATA CATATGCCAA GTGCCAAGTC AGGATCTGGC TGATACCCAG
 9301  GTAAGGAATA CCTGCATCGC GCAAGGCGCC GGTCAGACTG GCCGCAACCC
 9351  AGATCGTGGG CGTATCGGGC AGAAAGGCCG CGCCCGTTTC GTATGCGATC
 9401  CGCAGGGCCT GCAGGTTCTG CTCGGTGTCC CGCACCAGCA CGGCATCGTC
 9451  GCGCAACATT CCTTGCGCCA CCGGGATCTC CGACAGCACG ATGGTGTAGG
 9501  TATTGGCGGC CAGCGCTTCG GTGAAGCGCA ACTGGATGCC GGGAAACGGC
 9551  ACGCCCAGGT CGAAATAGAG CGCCCGCCGG ATCTGCAGCA GATCGTCGGT
 9601  GAGGGTGGCC GGCTCGAACC GGGGCTGCAG CCGCGCGGCT ACGTCGATGA
 9651  TCAGCGGGAC GGTGGGGGCG AATTCCGCCT GCCCATCCGC CGGCGCGCGG
 9701  GTGCGGGGCT GGCCGTCGGC AGCCATGCCG GCGAGCGCGG GCTCGGCGCC
 9751  TTCGGGCGGA CGCTGGGATG CGCGCAGCAG TACGAAACCG ATGGTGCCCA
 9801  CCGCGGCGGC CAGGGCGAAG AAGACCAGCG TGGGCATGCC GGGAATGAGG
 9851  CCCAGGCCTG CCGAGATCGC GCCGGCAATG ACCAGGGCGC GAGGCTGCGC
 9901  CAGCACTTGT GCGCCGATGT CGGTGCCTAC GTTGGAGGGG CCATCCCCGG
 9951  TCTGCACCCG CGTCACGATG ATTCCGGCGC AGATGGCGAT GAACAGCGCC
10001  GGGATCTGCG CGATGAGCCC GTCGCCTATG GTCAGGATGG CATATGTCTG
10051  CACGGCCTCG CCGGCGCTCA GGCCGCGCTG CAGCACGCCG ACCAGCATGC
10101  CGCCAAGCAG GTTGACGGCA ACGATGATCA GGCCGGCGAT GGCATCGCCC
10151  TTGACGAACT TCATCGCGCC GTCCATGGCG CCATACAGTT GGCTTTCCTT
10201  CTCGACCGTA CGGCGTCGGC GTCGGGCTTC GTCCATGTCT ATGGTGCCCG
10251  CGCGCAAGTC CGCGTCGATG GACATCTGCT TGCCGGGCAT GGCGTCCAGC
10301  GAGAAGCGCG CGGCGACTTC GGCCACCCGC TCCGCGCCTT TGGTGATGAC
10351  CACGAACTGC ACGATCGTGA GGATGAGGAA AACCACCAGG CCGACGATCA
10401  GGTTGCCGCC CACCACGAAG TTGCCGAAGG TCTCGATGAT GTGGCCGGCA
10451  TCGCCTTGCA GCAGGATCAG CCGCGTGGTC GCGATGGAGA TGCCCAGCCG
10501  GAACAGCGTG GTGACCAGCA GGACCGAAGG GAACGAGGAA AACGCCAGGG
10551  GCGAAGGCAG GTACATCGCG ACCATCAGCA GGACTGCCGA CAGCGTCATG
10601  TTCGCACCGA TCAGCACGTC GACCAGCGTT GTGGGCAACG GCAGGATCAT
```

Figure 5 (continued)

```
10651  CATGAAGACG ATCGCCACGA TGAGCACGGC CAGTACGATG TCGTTGCGGC
10701  TGGTGGCCAG CGCCACCGCG CGTTGCAGGC GGCGAATGGA TTTCTTGCTC
10751  GTCATGGCGG GGTCGCGCTC GCGCAAGACG CCGCCCGTAG CGCCATATAG
10801  TCACGCATGG CCTGGCGGGC GTCGTCGGGT CGGTTCAGCG CCTGCATGGC
10851  CTGCGCCCGG ACCAGATGGC CGGCGGCATC GGGTGTGGCG CGCAGTGCGC
10901  GCTTGTCCAG CGTGACCAAG GCCATGCGCG GTTCGCCCTG GTGCAGATAG
10951  CCCAGCGCCA GGGCCAGCAG GGACTGGCTG TCGATCGCGT CCAGGGCATC
11001  CAGGGCCGCC AGCAGGGCAA CCGTCTTGCT CCATTGGCGC TGCAACTGGT
11051  AGTGGTGCGC AAGCAATTGC AGAAGTTCGC GTACCTGAAG GCTGGGCGAG
11101  GGTAGGGTAT GCGGCATCAT CCCTGGTAGA GCGCGCTGCG ATACATGGCA
11151  GCCAGGTCGC GCAGCTTTCC GGCCTCGTTG AGTACGTGCA AGGCGCGGCT
11201  TAAGGCTGGC GCGGTGTTGC CGGCATGCAG TTCCATGGCC CGGGACAGTT
11251  CGTCGCGCGC CTGGGCCAGG GCGCGCTCGA ACTGCGCGGG CTGGAATAAT
11301  GCGCCATCCG TCAGGCGAGG CCGCGCCGAC TCGTCGAGCA TCGCGCTCAG
11351  GTCCGGGCGT ACGAGCAGGG CTTTCAGATG ATCGACCGCG CCGGTGGCGG
11401  GCGGTTCGAG CCAGCGCTCG GGTGGCAGGG TGGGGGCGGG CTCGCAGCGG
11451  GGACCGCGCA CGATGTGGTC GACGCCGCGC TCCAGGCCGA GGTGCATGGC
11501  ATGCATGCCG GGTACGGCGC TGGACATGGC GTCACGCCTC CAGGCGGTCG
11551  AGCCAGTTCG TCAGGCCCAG CACCGCCAGG CGCAACGCCG CTGCGTCGAG
11601  CGAGCGTTCG GGCAGTCGTG TCTGCGCGAC GATCCAGTCC TCGCTGCCCT
11651  CCGACCAGAG TGACGTCTGG ATGGATGCGG CGCTTCCGCG CTGCCCATGG
11701  GCCCGTTTCC ATGCCGCCAG CAACACGGAG GCGGCGTCGC GCTCGACCCG
11751  CTGGGCCAGG TGGACCAGGG CCGCGCCGGC GACGCATTCG ACGCCCAGGC
11801  GGCGCCCGTT GGACAGCGCC AGCGACGCCG ATCCGGACGG CCCGAATGCC
11851  AGGCCCTCGA TGCCGATATC CTGGCCGAAC TGATGCAGCG CCCTATCGGC
11901  AGTATTCATG CGTTCTCCAT TGCTATCGCG TTGTCCAGCG CGTCCTGCGC
11951  GGCCGCCAGG ACGGTGGCGC GCACGTCCAT GTCGGCGTAG ATCTGCGTGG
12001  GCAGGTCTTT GAGAATCTGG CGTACGCCAC CGAGGAATGC GATGCGCTCG
12051  GAGAGGGCAT TCGCACCGTG GCGCTCGGCC AGCTTCTCGA AGCGCGCGGG
12101  CGCAATCCAT TTGTCCTCGC TGATTCCCAC AAGATCGCGC ATCAGGCCCT
12151  GGGCGTCCGC ACACTCCTGC GAGCCTGCGT TGCCCAACCG TTGTTTCAGG
```

Figure 5 (continued)

```
12201  GCATTGCATT CCTCCAGTAC CGTGGCGGCC ACCTCGACTT GATAGAGATC
12251  GCTCGCCAAC ACTTGCAGCC TGACGCCGTC CGTCGACGGT GTCGCCGCGG
12301  CCAGGTCGTG TCCCAGCGCC TGAATCAGCG CGCCCAGCGC GCCGTGGATG
12351  TCGTCGTTCC CATAGCGTTC CAGCACCAGG TCCAGCGTGC GCGCCAACGA
12401  CAGCTGGCCC AGGGCGATGT CGCGGTACGC GTGCTGGAAG CCGGCCAGCT
12451  CGTCAGCGGA ACGCGCGAAT GCGCCGGCCG TGGGCAGGGT GTTGATGCCG
12501  GCGCGGATTT CGGGGCCATG GGCGAGCTCC AGGTCGGCCA ATGCATCGCG
12551  CAGGGCTTCG AGCGCGTGCG GCGCGGCGTC CTCGTGCTCG CCGCGCTGCA
12601  GCGCGTGCTG CAGCGCGAGG TATTGCTGCG TGACACCGGG AAACGCTTGC
12651  GCGGCCAGCT GCATGGGGGC GCCCCGGCCG CGCAGCAGCT CGGCGGTCAG
12701  GGCTTCCAGT TTTGCCTGCG CGTCGGGGTC GTGGGTGTGG GAAAACAGTT
12751  CCGCAAGCTG CGCCGCGTCC AGCCAGAGCA TCGGACGTTC GGCCGTGACC
12801  TTGCGTTCGG AGTGATGCTT TTCCTCGGCA GCCTGCGCCA TGTGCAGGCT
12851  GAGCTCCTCG GCCGCGTCCG CCAGCGATAT GCCGGTGGGC GCCGGTGCGA
12901  TGCGCTGGCC TTGCAGCCAG CCGGAGGAGG TGTTGGCCGA GGCGTCGTGG
12951  CGCCCCTGCA TGGCGGCGTG GAAGGGATTG GGGGCGGCAT CGATACGAGT
13001  CATGGGGAGT CCTCGGAGAA GGAACCATTT GCCTACTGGT GCAGTGAGTG
13051  TCGCGCGCGC GGTCATGGTT CCCGGAAACG GGCGCGATAT TGGGCAATTC
13101  GCAGCCTGGA ACTTGCCGCG GGCGCAGGGT TACTCAGCAT GCGTCTTTCA
13151  ACTCGAAGGA GCTCTCATGA GCATTGATCT CGGAGTTTCA CTCACGTCGC
13201  AGGCCGGCGG CCTGCAAGGC ATCGACCTCA AGAGCATGGA TATCCAGACT
13251  CTCATGGTGT ATGTGCAGGG TCGTCGCGCC GAACTCCTCA CGGCTCAAAT
13301  GCAGACCCAG GCCGAAGTGG TGCAGAAGGC CAATGAACGC ATGGCGCAGC
13351  TCAACGAGGT CCTGTCCGCG CTGTCCCGGG CCAAGGCCGA GTTTCCGCCC
13401  AATCCGAAGC CGGGCGACAC CATCCCGGGC TGGGACAGCC AGAAGATCAG
13451  CCGGATCGAG GTTCCTCTCA ATGATGCGCT GCGTGCCGCC GGCCTGACGG
13501  GCATGTTCGA AGCGCGCGAT GGCCGGGTGA CCGGCCCCGA CGGCCGGGGT
13551  ACGCAGGTCG TGAACGGCAC GGGCGTCATG GCCGGTTCCA CGACCTATAA
13601  GGAACTCGAA AGTGCCTACA CCACCGTAAA GGGGATGCTG GATACGGCGT
13651  CCAATACGCA ACAGATGGAC ATGATCAGGC TGCAGGCCGC CAGCAACAAG
```

Figure 5 (continued)

```
13701  CGCAACGAGG CTTTCGAGGT CATGACCAAC ACCGAGAAGC GGCGCAGCGA
13751  CTTGAACAGC TCCATCACCA GCAACATGCG CTAAGCGCTG CACAAGGAGT
13801  ATTCCATGCA GGAGCAAGGC ATCCAATCCA TCATGCGCGC CGCGGAAGAG
13851  CTGGTCGAGC AGACCCGCCA GGCGTTGTAC AGCGTCGACG AGATCTACGC
13901  CCACGTTGGC GTCGACCCCG CTCGCCTGCG CAATCTGGCG GTCGAGCAGG
13951  CCAGGATAGA GGCCGAGGCC CAGGCGGCGT TCCGTGATGA CCTCGCGGAC
14001  ATCGAGCGCG AGGCGGCGCG CGTCAAGGCG GCCTGCACCG ATGCGCCGCA
14051  GGCCCGCAGG GTGCTTCACA ACCACGTCTG AGCGCGGAGG CCTTCCATGC
14101  CAAAGTCAGC CGACCAGGGC GGCTCCCCGG CGTCAGCTTC GCATGAGGCG
14151  TTGCGCCATA TTCTCGACGC AGGCGCTTCG ATGGGGGGCT TGCAGGGGTT
14201  GGACGAGGCG CAGCAGCAGG CGTTGTACGC GATCGGTCAT GGCGCCTACG
14251  AACAGGGGCG CTATGCCGAC GCGTTGAAAA TGTTCTGCCT GCTGGTCGCG
14301  TGCGATCCGC TGGAAGCCCG TTATCTGCTG GCCCTGGGCG CCGCGGCCCA
14351  GGAGCTGGGG CTGTACGAGC ATGCCTTGCA GCAATACGCG GCCGCGGCGG
14401  CTTTGCAGTT GGACTCCCCC AGGCCCCTGT TGCATGGCGC CGAGTGCCTG
14451  TATGCGTTGG GTCGTCGCCG CGACGCCCTG GATACGCTCG ACATGGTGCT
14501  TGAGTTGTGC GGCTCGCCGG AGCGTGCGGC CCTGCGCGAA CGGGCCGAGT
14551  TGCTGCGCAG GAGCTATGCA CGTGCCGACT GAAACGGCGC CATGTCCGCC
14601  GTCAAGATTT CAATTCGAGG AGGTTCGATA TGTCTGTTTC TCCGACTTCG
14651  CCCGGCTCTT TCGGGGCCGG CCCTGTCTTT GACTCCGAAT TGCAGGCCCC
14701  GGCCCCGTCG GCGCAGCGTC GCGGCGGTGC GGCGCCTGTG CCGCCGCCCG
14751  TCGATCGGCG CGGCGTCGAG CCGGGAGATC CCACGCTGGG CATGCTGCCC
14801  GCGCCAGATT TGCTCGCGGG GGGCGCCGTC AGCCGCACCC GCGCGGCGCT
14851  CGACGATCTG GACGCCGCAC GGCTCGGTGA AGACATCTAC GCCTTGATGG
14901  CGGTGTTGCA ACAGGCCAGT CAGCAGATGC GGGACGCCGC CCGTATCGCT
14951  CGTGATGCCG AGGCTACGCG GCAAACGCAG GCTCTCGGCG ATGCGGCCAG
15001  CCAGATGCGC CAGGCGGCGA GCGAGCGCAT GGCCGGAGCG ATCGTGGCGG
15051  GCGCCATGCA GATAGCGGGT GGTTTCGTGC AGCTGGGGGC GGGCCTGGCA
15101  GCGGGTTTGC AGGCCATGGG TGGCGCTGCT GCGCAAGCCA AGGGCGCCGC
15151  ATTCTCCGAG CAGGCCTCGA CAAGCCGCAA GGTGGCGGCC GGCTTGCACG
15201  ATGCCCCCGA GCTGCAGGCA ACGGTGCAGG CCCGCGCAAC CCAGCTCGAA
```

Figure 5 (continued)

15251 GCGCAAGCGG CCTCGTTTGG TGCGGACGCG GCTCGTTCGT CGGCAAAGTC
15301 GCAGCGCGTA TCGAGCGTTG CCCAGGCCGG CGCCGCAGCG GCCGGCGGTA
15351 TCGGCGGCCT GACCAGCGCC GCCCAGGAAC GCCGCGCCGC CGAGCACGAG
15401 GCCAGGCGCG CGGAGCTGGA CGTCGAAGCG AAGGTGCATG AAACGGCCTC
15451 GCGGCGGGCC GACGAAGCCA TGCAGCAGAT GCTCGACATC ATCCGCGGCA
15501 TCAGGGAAAA GCTGGCCGGG ATGGAGCAGT CCCGCAGCGA GACCGCCCGT
15551 AGCGTGGCCC GCAATATCTG AGTGTCCGGC TCCAACCTTC AATCTTGAGG
15601 ATGACCGTCA TGAGTACGAC CATATCCACA GCCCCGAGCG GCGCCGCGCT
15651 TGCGCCGTCT CGCATAGATA TGCGGGCGCC GGAGCCCGGG AGTGCCGGCG
15701 AAGGCGCCGG TATCCTGGCG CCGGTGACGA CGCTGGCTCT GGCGGCGGGC
15751 CGGCCGGCTT TGCCAGCGTC ACCGTCGCTG CGCACCGCGC CCGTCCTGGA
15801 TCCGCCAGTG CGCGATCTCA GCCCCGCCGA CTTGGCCGAC CTGCTGCGCG
15851 TCTTGCGATC CAGGGCGGTG GACGGGCAGT TGGCCACGGC GCGCGAGAAC
15901 CTGCAGGATG CGCAAGTCAA GGCGAAGCAG AACACCCAGG CCCAGCTCGA
15951 CAAGCTGGAC GCATGGTTTC GGAAGGCTGA GGACGCCGAG AGCAAGGGCT
16001 GGCTGAGCAA GGTGTTCGGC TGGATCGGGA AGGTGCTGGC GGTCGTGGCA
16051 TCGGCCCTGG CTGTGGGCTT TGCTGCCGTC GCCAGCGTGG TCACCGGCGC
16101 GGCGGCCACG CCCATGCTGG TGCTCAGCGG CATGGCATTG GTCAGCGCCG
16151 TGACATCGCT GGCCGACCAG ATATCGCGAG AGGCGGGAGG GCCGCCTATC
16201 AGCCTGGGCG GGTTTCTCTC CGGGCTGGCC GGACGTCTGC TGACAGCGTT
16251 GGGGGTGGAT CAGTCGCAGG CCGACCAAAT TGCCAAGATC GTCGCCGGCC
16301 TGGCCGTGCC CGCCGTCTTG CTGATCGAAC CCCAGATGCT GGGCGAAATG
16351 GCCGAAGGCG TGGCCAGGCT GGCGGGCGCC GGCGATGCCA CCGCGGGATA
16401 CATAGCCATG GCGATGTCCA TCGTGGCGGC GATCGCGGTC GCCGCGATCA
16451 ATGCCGCCGG TACGGCCGGC GCGGGCAGCG CCTCGGCGAT CAGGGGTGCC
16501 TGGGATCGGG CCGCCGCGGT AGCCACCCAG GTCCTTCAGG GGGGTACGGC
16551 AGTGGCGCAA GGCGGCGTCG GCGTGTCGAT GGCAGTCGAT CGCAAACAGG
16601 CCGATCTCCT GGTCGCCGAC AAGGCGGATC TGGCGGCGAG CCTGACAAAA
16651 CTGCGGGCGG CCATGGAGCG TGAGGCGGAC GATATCAAGA AGATCCTGGC
16701 TCAATTCGAC GCGGCCTATC ACATGATCGC GCAGATGATC AGCGACATGG

Figure 5 (continued)

16751 CGAGCACGCA CAGCCAGGTC AGCGCCAACC TCGGACGGCG CCAGGCGGTG

16801 TAGCGCCGGG CGCTCAAGGA ATTTTCATGA CTGTTCATGA CGACGCGCCC

16851 GCGGCGCTGC GCGCCCGGCT GGATGCGTTG CCGGGCAGCC GGCGCCTGAC

16901 AGCCGAGCAA TTGGAAGTGA TTTACGCGAT GGCGTATGCG CACGTCGCCA

16951 GGTGCGAGTA CGGCAAGGCG CTGCCCATTT TCGCCTTCCT CGCGCAGTAC

17001 GGCCCCACGC GCAAGCACTA CTGGGCCGGC CTGGCGCTAT GCCTGCAGAA

17051 GACCGACCGT CCCGACGAGG CGCGCAATAT CTATGCGTTG ATCCTCACGC

17101 TCTATCCAGA TTCCGCGGAT GCCGTGTTGC GCACGGCCGA GTGCGAGCTG

17151 GCGTTGGGTG AGAACGAACG GGCGCAGGCG GCCCTGTTCG GCGCAATCGC

17201 CATCGATGCA GAAAGTGGGC AGCCAGGTCC GGTCTCGCAC CGTGCGCGCG

17251 CTTTGCTCGA TCTTATTTCA GTTTCACATC CGGAGTAACT CCATGCACTC

17301 AGACTCAGGT TCAGATTCAG GCTCAGACTC AGGCTCAGGC TCACCCATGG

17351 TCTCGTCGAT ACATCCATCG GAACCGATAC AGCCGATGGA GCATGTGCTC

17401 GAGGAGGCCG ACGCCCGCCT GCTTACCGAA GTGGGTTTTC TGGCGGCGGC

17451 CGTCAGCGAT CTGACGCGCG CGGACGCCAT TTTCAATGCA TTGCAACGTG

17501 TACGGCCGGG CCGGACGCAT CCCTGCATCG GCCTGGCGGT CGCCCGCATG

17551 AACGCCGGGC TGCCCGACGA AGCCGCCGAG ATCCTGGCGA ATTTCCAGCC

17601 GGCACAGCCG GAGGACCGCT CGGAACTGGA CGCCTGGTGC GGGTTCGCTC

17651 TGTTGCTGGC TGGCCGCTCG GACGAGGCGC GCCGCATGCT GCAGCGAGCC

17701 ATCGATGCGG GTGGCGAGGC GGCAAGGCTG GCGCAGGTCG TGTTGGACAG

17751 CGGACCCGCC ATGATGCGGC CCGCGCCGTT GCAGTCCGAG CCATTACCTG

17801 GAGCTCCTGG ATGAATTTGG ATCTGACGGC GATCAACGCC GTGCAGGAAC

17851 GGCTGCTCGC TGGATCATTC GACATGCCGC GATCTCCCGC GATGGCGGAT

17901 CAGGCGCGCT TTGAATTGGC GCTGGGCGAG ATGCCCGGCG CATCGGCCCC

17951 GAACGGGGCG ATCGCCCTGG CGCCGGTCGC GCTCGACGAG CCGCTGGGCC

18001 GTCGCATTCT TGGACAGTTG CGCGGCGGCC TGGCCGATGT GGCAGGAAAA

18051 TGGCGGGCGG TGCAGACGGG CTTGGCCGAG GTGAGCCAGG CGCCTACCGT

18101 GGTGGGTATG CTCGATCTGC AGGCCAGGTT GCTACAGGCA TCCGTGGAGT

18151 ACGAGTTGGT GGGCAAGGCA ATAGGGCGCG CCACCCAAAA CGTCGATACG

18201 CTGGCGAGAA TGTCATGAAC GCCATCGGGG CGATCCAACG GTATCGGCGC

18251 GGCGCGGGAT GGGCGGCCCT GGTGCTCGCC CTGGCGCTGC TGGCCGGCTG

Figure 5 (continued)

18301 CGGTGCCCGC GTCGAGCTGT TGGGCGCGGC GCCCGAGAAC GAAGCCAACG

18351 AAGTATTGGC GGCGCTGCTC GAGGCAGGCA TCGCTGCGCA GAAGCAGTCC

18401 GGCAAGGCCG GCTACGCGGT TTCGGTGCCG GCCGAGGCGG TGGCCCGGTC

18451 GCTGGAGATC CTGCGCGCAA GCGGCCTGCC CCGCGAGCAG TTCGACGGAA

18501 TGGGACGCAT ATTCCGCAAG GAAGGCCTGG TTTCATCTCC GCTCGAAGAG

18551 CGCGCCCGCT ACATTTATGC GCTGTCTCAG GAATTGGCCG ACACCCTGTC

18601 GCAGATCGAC GGCGTGCTCA GCGCCCGCGT GCACGTGGTG CTTCCCGAAC

18651 GCGGCGCGGT CGGCGAGCCG GCCACCCCTT CGACGGCAGG GGTGTTTCTC

18701 AAGTACCGCG ACGGACAGAG CCTCGACGCG CTCGTGCCCG AGATCCGCAA

18751 GCTGGTCACG CATGCCATCC CGGGCCTGGC CGAGGACCGT GTATCGGTTG

18801 CCCTGGTGGT GGCCCAGCCC GTTCAGGCCG CACCCGCGCC GGTCGCGTGG

18851 CGCCGCGTGC TTGGCGTACA GGTCGCGGAC GGATCGGTCC TGAGATTTTC

18901 GCTGTTGCTG CTGTTGTTGC CGGTGCTGTG CCTGATAGTG GCGGGGGCCG

18951 CGCTCTACGT CTGGCGCACG CGCTGGTCCC GCGGCGAAGG GCGCGGCGGC

19001 GCTGGCGCCG GCGCCACGGA AGGAGCCGGG CATGACTGAG GCGAGCGTGC

19051 TGCTTTCCGA GCGGCTCATG ATATTCAATC TCCTGCCCAG CCTGACCCTG

19101 CATGCCAGTC GCCACGACGA GATGTTTCCA GCCGATTGGG TGCGCGCGTT

19151 GTGCAATGCC GACGCGGCGT TGGCCAACGC GTGGCATCGC CATTGGTCGC

19201 GCTGGATCTT GTGCGAATTG GGCCTGCTGA ACCAGCCGGT CCTGAGCCTC

19251 GATCCGCCGC AGTTGAAGGT CGCGCTATTG TCCACGGACG CCTTGCGGAC

19301 CTGCGCCGCC CATGCGGGAG CGCTGCTGTG CGCGCCGCGC CTGCGACGCG

19351 CGATAGACGG CGCCGAGGTC CGTACCTTGC ATGCCGCGCT CGGGCGCGAT

19401 GTGATGAATT TCGCCGTGTC TTCCGCGGCG CGGGCCCTGC ATGACGGGCT

19451 CGCCGCCAGT TCGGACTGGA CCCTGGCCGC CACGGTCCAG GCGGCGCAGA

19501 AACTGGGCTG GGCCCTGCTG CGCGACGCCG TGCAGGGCGC CGCCGACGAG

19551 ATAGCGCTGC GTTGCGCGCT GAAGTTGCCG CGCGACCTTG ATCCCGCGCC

19601 CGTCCTGCCG CCCGAGGCGG CGCTTGCGCT GGTGCTGTCC ATGCTCGAAA

19651 TCCTGGATGC AGAATGGCTT TCCTCGTTCC CCGCCCAAGC CTGATCCAGG

19701 CGGTACGGCC CGGCCGTGCG GATCCCGCGA CCGACGTCTT GCGCGCTGAA

19751 GACTACGCCG AGCTGCTCAG CGCCGCGCAG ATCGTTGCCC AGGCACACCG

Figure 5 (continued)

```
19801  GCGGGCCGGC GAAATCGTGG CCGAGGCGCG AGAGGAGTTC GAGCGCGAGC
19851  GCAGGCGAGG CTATGAGGAG GGGCGCCGCG AAGCGCTTAC GGATCAGGCG
19901  GAGAAGATGA TAGAAACCGT AAGCCGCACG ATCGACTACT TCGCGGGTAT
19951  CGAGAACGAG ATGATCGAAC TGGTCATGAG TGCGGTCCGC AAGATCGTCG
20001  ACGGTTACGA CGACCGCGAG CGCACCGTGA TCGCCGTGCG CAACGCATTG
20051  GCGGTCGTGC GCAATCAGCG CCAGATGACC TTGCGCCTGC ACCCAGACGA
20101  GGTGGATGTG CTCCGGGAAG GCATGAACCA GCTTCTGGCG GCCTATCCGG
20151  GCGTGGGCTA CCTGGACCTG CTGCCCGACG CCAGGCTGGC GCCGGGAGCC
20201  TGCATACTGG AGAGCGAGAT AGGCATGGTC GAGGCCAGCC TCGAGGACCA
20251  GCTGTGCGCC TTGCGGGCGG CCTTCGAACG TACATTCGGC CGGCGCGGAT
20301  AGGGGCATGC GTCAGTACCA CTACATCACG GAGATGATGC GGGTGGCCCT
20351  GCAGGATCTG TCCACGCTGC GGATAAAGGG CCGGGTGGTG CAAGTGGTGG
20401  GAACGATCAT CAAGGCCGTC GTTCCGATGG TCAAGATCGG CGAAGTGTGC
20451  CTGCTGCGCA ATCCCGGCGA GGACTTCGAG ATGCACGGCG AAGTGGTGGG
20501  CTTTGTCCGC GACGCCGCCT TGCTCACGCC TATCGGCGAC ATGTACGGGA
20551  TTTCCTCGGC GACCGAGGTG ATACCGACCG GACGCACGCA TATGGTCCCC
20601  GTCGGTCCGG GCTTGCTGGG ACGCGTGCTG GACGGGCTGG GACGTCCGCT
20651  GGACGCCGCC GAGTCAGGGC CGCTGCATGC CCACAAGTTC TATCCGGTCT
20701  TCGCCGATGC GCCAGACCCG CTGACGCGTC GCATCATCCA TGCTCCGCTG
20751  GAGCTGGGGG TGCGCGTACT GGACGGTTTG CTTACATGCG GGGAAGGCCA
20801  GCGTCTGGGA ATTTTCGCAG CCGCCGGCGG CGGCAAGTCG ACCCTGCTGG
20851  GCATGCTGGT CAAGGGCGCC GCGGTCGACG TGACGGTGGT GGCGCTGATC
20901  GGCGAGCGTG GGCGGGAAGT TCGCGAGTTC CTTGAGCACG AACTCGGTCC
20951  GGAGGGCAGA CGCAAGAGCG TGATCGTCTG CGCGACCAGC GACAAGTCCT
21001  CGATGGAGCG TGCCAAGGCG GCGTACGTCG CAACCGCCAT CGCCGAATAC
21051  TTCCGCGATC AAGGGCAGCG TGTACTTTTT CTGATGGACT CGGTCACCCG
21101  CTTTGCGCGA GCCCAGCGTG AAATCGGCTT GGCGGCAGGC GAGCCGCCGA
21151  CGCGGCGCGG CTATCCACCG TCGGTGTTCG CCACCTTGCC CAAACTGATG
21201  GAGCGCGCCG GCATGAACCA GACGGGTTCG ATCACGGCGC TGTATACGGT
21251  GCTGGTCGAG GGGGACGACA TGAACGAACC GGTGGCCGAC GAGACGCGTT
21301  CGATACTGGA CGGCCACATC GTGCTCTCGC GCAAGCTGGG AGCGGCGAAT
```

Figure 5 (continued)

```
21351  CACTATCCTG CCGTCGACGT GCTGGCCTCG GCCAGCCGGG TCATGAATGC
21401  CGTGGTGTCG CCGCGTCACA AGTACCTGGC CGGACGTATG CGCGAACTGA
21451  TGGCCAAGTA CCAGGATGTC GAGCTGTTGG TGAAAATCGG CGAGTACAAG
21501  CAGGGCGCCG ATGCGTCGAC CGATGAGGCG ATACAGAAGA TCGGACAGAT
21551  CAATGCGTTT CTCAGACAAC TAACCGACGA ACGCGAAGCA TTCGAGGATA
21601  CCGTACTGCG CATGGCTGAA ATCATCGGAC CCGAATCCTA ATGGACCTGG
21651  AAAGCCTGCT TGCCATCAAG CATTTTCGCG CCGACCAAGC CCAGCTTGCG
21701  CTGAAACGCC AACAGCAGGC CTGCGCGGTT GCTGCCGCGG CGCAGCGTCA
21751  GGCGCAAGGC CGCCTCGACG ATTGTCGCCT GTGGGCCGGA CAGCTCGAAA
21801  ACCGTCTATA TGCCGAGCTG TGCCGGCGCA TCGTCAAGAC ACGCGACATC
21851  GACGAGGTGC TGCAACGAGT GGGCCACGCC CGCGACCGCC AGGCCAGCCT
21901  GGCGCTGCAG CTCGACGACG CCGTGCGCCG TCACGAACAT GAAATCCAGC
21951  TGCTCGCGCA GCAGCGCGAG CAGCACCGGG AGTGCTTCCA GGCGCAGCAA
22001  CGGATCGCCG AGTTGGTGCG CCTGCAGCAG GTCGAGGCGG CGGCCTTGCG
22051  CGAGAGCCAG GAAGATCGCG AAATTCAGGA AGCCATCGAA TTGTCGGCGC
22101  GTGGGCGCGA CGATGCATCG CGAGCCGGCG ACGGCCTGGC GCGGCTATGA
22151  ACCAGCCAGA CGGGCTGGGT TCGCCCATGG CCGGCGGCGG GCAGCGCATG
22201  GGCGTGGCGC GCACGCCGTA TGCGCGTCAG CCGGATCGGG ATGCGCAGCG
22251  TGCCTTCGAG CGGGAAATGG AACAGGAGAA AGCGAAGGAA GAACTGCCCG
22301  GGCCGCAACG CCTGGCGCCG GGTCCGGCCT GCGTCGGCTG GCTGGCGTCG
22351  ATGGAACCTG CCGCCGGCCG TCCACCGGCC AGTCTGGCCC AGGCGCTGGC
22401  AAGCGTGGCT GCGGGGCTGG CGGTAGGCGA CGTGCTGGAG GGGTATCGCG
22451  AAGCCCGTAT CGTTGTGGAC GATACGCTGC TACCCGACAC CACCTTGTCG
22501  GTACGGGAGG ACGGCGGCTG GATCGTGGTG GCTTTCGCAT GCCGACAACG
22551  GGACGCTTGC GAGCGCCTGC ACGCGTGCGC CGACCGGTTG GCCATGGAGC
22601  TCGCGCTGGA GCTGGCGCGC GACGTCGAGG TTGCGGTGGC ATGCGACGGC
22651  GAGCCGCACG AGCGGGTGGC GCGCGCGCAG CGGCCGTGGC GATGAATCGA
22701  GTGGCCGGCG GGGCGGCGGC GCAGGCCGCT GGCATGGTGG ATCTCGCGGT
22751  TCCGCGGTTG AGCGCCGGCG AGGCCCATGC CCTGTCGAGG ATTGCATGCC
22801  ATGGCGCGCG ATTCGACGTT CGGCTTGGCG AGCCGGCCGT GCGCTGGCAC
```

Figure 5 (continued)

```
22851  TGCGCCCTGA CGCCTTGCGT GCACGGCGAC CTTGCCGATG GCGAGATGGA
22901  AAGCCTGCAA CTGCAATGGG CCGGGACGTA CATCGGCCTG ACGGTTCCGC
22951  GCGCGGCCGC GGCGGGATGG CTGGCGGCGC GCCTGCCCCG GTTTTCCGGC
23001  GTGGAGTTGC CGGAACCCAT TGCGGCGGCG GCCCTGGAGG CAATGCTGGA
23051  GGAGGTCTGT CGAGGCGTGG CCGGACTCGA CCAGCAAGGC CCGGTCCGCG
23101  TGGCGCGGCA AGGCGGGACG CCACCGGTCC AGCCGCATCG CTGGACCCTG
23151  ACGGTACGGG CGCCTGACGG TGGCGTCTGG CGCGCGGTAC TGGCGTGCGA
23201  CGCATGGGCC TTGCAAGCGG TCGCGGCGGC GCTGGATTCC GTTGCGCCTG
23251  CCGATGGTCG GGTCAATCCG GAGCGCGTGC CGGTCAGGTT GCGTGCCGAT
23301  GTCGGCGCGG CGTCCGTGAC CGCAGGCCAG CTGCGGACGC TGCGAGCGGG
23351  CGACGTCGTG TTGCTCGCGC AGTACCGGGT GAGCGATGCC GCAGAACTAT
23401  GGTTGTCGGC CGGACCCAGC GCGATCCGGG TACGGGCCGA GCATGCGTCT
23451  TTTCGTGTAA CTCAAGGTTG GACTCCCATC ATGACGGAAC CCGCGACACC
23501  TGACCCTGGC GAAACCCCGG CACAGGCCGA CGCGACGCTC GATACCGATC
23551  AGATACCCGT GCGCCTGACG TTCGACCTGG GCGAGCGCGA GTTCACGCTT
23601  GCGCAGCTGC GCAGCCTGCA TCCGGGCTGC ACGTTCGACC TCGAGCGGCC
23651  CATCGCCGAC GGGCCGGTCA TGGTGCGGGC CAATGGCCTG TTGCTGGGCA
23701  GCGGCCGGCT GGTCGACATC GACGGCCGCA TCGGCGTGGT ATTGCAGTCG
23751  GTCAGGCCTG GACTCGCATG AGCGATACCG ACCCCTTCAG CCTGGCCCTG
23801  TTTCTGGCGC TGCTGGCGCT GGTACCGCTC ATCGTCGTCA TGACCACGTC
23851  GTTCCTGAAG ATCGCCGTCG TGCTTGCCTT GGTGCGCAAC GCCCTGGGAG
23901  TGCAACAGGT ACCGCCCAAC ATGGCCCTGT ACGGGCTGGC GCTTATTCTT
23951  TCCGCGTATG TGATGGCGCC GGTCGTTCAC AGGATAGGCA CCGAGGTCCA
24001  GGCCTTGACC GCGCAAGCCG GGGAGTCCGG CACCGCCGCG CCGATGGCGC
24051  TGGACGCCGT GCTTGGCGTG GCCGAGCGAG GCGTGGGGCC GCTGCGGGCC
24101  TTCATGTTGC GCAACAGCCA GCCGGCCCAG CGTGATTTCT TCCTGCGCAC
24151  AGCGCGTCAT CTCTGGGGCG AGGAGGCATC GCGGGACCTG TCGGAAGACA
24201  ACCTGCTGGT ATTGACGCCC GCATTTCTGG TTTCGGAGCT GACCGCCGCA
24251  TTCCAGCTTG GCTTTCTGCT GTACCTGCCG TTCATCATCA TCGACCTCAT
24301  CGTATCGAAC ATTCTTCTTG CCATGGGAAT GATGATGGTT TCTCCCGTGA
24351  CGATCTCCAT GCCGTTGAAG CTGTTCCTGT TCGTCATGGT GGACGGCTGG
```

Figure 5 (continued)

```
24401  ACGCGCCTGA TCCAGGGCCT GGTGCTTTCC TATCGGTGAC CAGCATGCAA
24451  ACCCAAGACC TGGTTTCGTT CATGACACAG GCGTTGTACC TGGTGCTCTG
24501  GCTGTCGCTG CCGCCCATCG CCGTGGTGGC GATCGTGGGA ACGCTGTTTT
24551  CCCTGTTGCA GGCCTTGACG CAGGTGCAGG AGCAGACCCT GTCCTTCGCC
24601  GTGAAGCTGA TAGCCGTGTT CGCCACGCTG ATGCTGGCGG CCCGGTGGAT
24651  AAGCGCGGAA ATCTATAACT TCACGATTGC GGTGTTCGAT GCCTTTCATC
24701  GGATCCACTG AGCGGCCAAT CGATGCACAC GGAGTTCAAT TTCGTCGAGG
24751  CGAAGGTTTT CCTGGGAACG CTGGCCATGA CGCAACCGCG GATACTCACG
24801  GCCATGCTCT TTCTGCCGAT GTTCAACCGT CAGTTTCTGC CTGGTCCGCT
24851  GCGTTACGCC GTCGGCGCCT GTCTCGGGCT GATCGTGGTT CCCCAGCTGG
24901  CGCCGCAGTA TGCCGCGCTG GATATCGACT GGCCCCGGCT GCTGGCGCTG
24951  CTGGCCAAGG AGGCGATGGT GGGCATGTTC CTGGGTTGGC TGGCTGCCTT
25001  GCCATTCTGG ATCTTCGAGG CCATCGGCTT CGTCATAGAC AACCAACGGG
25051  GCGCCAGCCT GGGCGCTATC CTCAACCCCG CCACGGGCAA CGATTCGTCG
25101  CCCATGGGCA TTCTCTTCAA TCTGGGATTC ATGGTGTTCT TCCTGACGGC
25151  GGGCGGATTC GGGTTGTTCG CCACGATGCT GTATGACAGC TTCGGGTTGT
25201  GGAACATCTG GGCGTGGTGG CCGTCCATGC CCGCACAGGG CGCCGTGCGG
25251  ATGCTGGACC AGTTCAGTGG CTTTGCCGCG CGTGTCCTGC TGCTGGCCTC
25301  GCCGGCCATC GTGGCCATGT TCCTGGCCGA GCTGGGCCTG GCCCTGATCA
25351  GCCGCTTCGC GCCTCAACTG CAGGTGTTCT TCCTGGCTCT GCCGGTAAAG
25401  AGCGCGCTGG TGCTGTTCGT GCTGGTGCTG TACATGGCAA CGTTGTTCCA
25451  GTATGCAGGC GAAATCCTGG GTTCTGTGGG CCGGATCGTG CCGTTCCTGC
25501  ATTCAGCGTG GCCCGGCCCA TGAGCGGCGA GAAAACCGAG CGGCCCACCC
25551  CGAAGCGCCT GCGCGATTCC CGCGAGAAAG GCGAGGTCGC ACACAGCCGG
25601  GACTTTACCC AGACGGCGCT GATATGCGCC TTGTTCGGGC ACTTTCTGAT
25651  CAATGCCCCG TCCATTCTCG CGTCGCTGCG AGCGCTGATA CTGGCGCCGG
25701  CGGCCTTTGC CGACCAGGGG TTCGCCGTCG CATTGGGGCC CGTGCTGACG
25751  GAAATCCTCG ATCAGGCCGT CCGCGTGCTC GCTCCGCTGA TTCTCATCGT
25801  GCTTGGGGTG GGGATGTTCG CCGAATTCCT GCAGGTAGGC GTCGTGCTGG
25851  CGTTTCGAAA GCTCAAGCCT TCGGCGGAGA AACTGAATCC CGCCGGCAAT
```

Figure 5 (continued)

```
25901  TTGAAGAATA TCTTCTCGGC GCGCAACCTG ATGGAGTTCA TCAAGTCGGT
25951  ATGCAAGATC CTGTTTCTGG CGGTGTTGGT CACGTTGGTG ATACGGGATT
26001  CCTTGCAGCC GCTGATGGCC GTTCCCCATA GCGGGCTGGA CGGGTTGCGA
26051  ACGGGCGTAG GCCGCATTCT GCAGGTCATG GTCTGGAACA TCGGACTGGC
26101  GTACGGGGCG ATTTCGCTGG CGGACCTGGC CTGGCAGCGT TACCAGTATC
26151  GCAAAGGCTT GCGGATGAGC AAGGACGAAG TGAAGCAGGA GTACAAGGAG
26201  ATGGAAGGCG ATCCCCATAT CAAGCAGCAA CGCAAGCACC TGCACCAGGA
26251  GCTGATCATG CATGGCGCGG CGGCCCAGGT TCGCCGGGCG ACGGTGCTGG
26301  TGACCAATCC GACACACCTG GCCGTGGCCC TGTACTACGC GGCGGGCGAG
26351  ACGCCCTTGC CGCGCGTGCT GGCCATGGGG CAGGGAGCCG TGGCCGCTCT
26401  CATGGTCGAG GCCGCGCGCG ATGCCGGCGT GCCGGTCATG CAGAACGTCG
26451  CGCTGGCCCG CGCCTTGCAC GACCAGGCGG AGGTGGACCA ATACATTCCC
26501  GGCGAGTTGG TGGAGCCGGT GGCCGCGGTG TTGCGGGCGG TGCGCCAGGC
26551  ACTCAAGGAG CAGACATGAC AGCAACCATT CATCCCGATA TTGCCGATTA
26601  TGCGCGACGC CATGGCCTCG AACCCTCGGT CGACGCCGAT GGCGGGCTTG
26651  CCGTCCGGAT CGACGGACGG CATCGCGTCA GGTTGATCCC CGCCGAAGAC
26701  GGCATGCTGG TGTTGCGGGC GCGGCTGGCC GAGCTGCCCG ATGGGTGGCA
26751  GGCGCGCGCG GCGCAGTTGC GCCGGGCGGG CCTGCTGGCC AGCGCCATGG
26801  CCCCTGCGAC CGATGCGTAC TGCGGCATAG ACCAGGGCGA AACCGCGTTG
26851  TATCTGCACC AGCGCGTCGC ACCGGCCGGC AGTGCGCTGG CGGTGGACGA
26901  GGCGGTGGGC GAGTTCGTCA ATGCCTTGGC CACTTGGAAA AGGGCGATGG
26951  CGCAATGGCA ATAGGTCGGC TTGGGTATCT TGTCCGCGGC GCATGGGCCG
27001  GGGGTGTCAT GCTGTTGGCG GCCGGTAGCG CCTGGGCGGC GCCGAACTGG
27051  CCTTTGGCGC CGTATAGCTA CTACGCGCAG CAGCAGAGCC TGTCCGATGT
27101  GCTGCGCGAG TTCGCCGCAG GCTTCAGCCT GGCGTTGCAA CAGGGCAAAG
27151  GGGTGCAAGG CGTGGTCAAT GGGCGTTTCA ATGCGCGCAC ACCCACGGAG
27201  TTCATCGAGC GTCTCAGCGG CATCTATGGG TTCAACTGGT TCGTGCATGC
27251  CGGCACGCTG TATGTCAGCC GCACCAGCGA CGTGGTTACC CGCGCGGTGG
27301  ATGCAGCCGG CGCTTCGCCG TCGGCGTTGC GCCAGGCCTT GCTGCAACTG
27351  GGCATCCTGG ACGAACGCTT CGGATGGGGA GAGCTGCCGG CGCAAGGCGT
27401  GGCCATGGTG TCAGGGCCGC CGGCCTATGT CGCGCTGGTC GAGCAGGCGG
```

Figure 5 (continued)

```
27451  TAGCGGCGTT GCCCAAGGGG GCCGGCAATC AGCAGGTGGC GGTGTTTCGC
27501  CTCAAGCATG CTTCCGTGAG CGACCGGGTG ATCCGTTATC GAGACCAGCA
27551  GGTAGTTACG CCGGGGATGG CCACCATGCT GCGCCAATTG ATCCTGGGGG
27601  CGGGGCCGGG CAACGACGCG GCGCTGGCCG CGGTGGCGGC GCCGCTGCGG
27651  GAAAATCCGC CGGTGTTCGG CGATGCGGCA GCTGACGGGA ACGCGCCGCT
27701  CGCTGGCGCA GCCCAGGCAG CCGGCCGGCG CCTGAGCGAG CCCAGCGTGC
27751  AGGCCGACAC GCGCCTCAAT GCCTTGATCG TGCAGGATAT TCCCGAACGG
27801  ATGCCAATCT ACCGTGCCCT GATCGAGCAG TTGGATGTGC CCAGCACCCT
27851  GATCGAAATA GAGGCCATGA TCGTGGACGT CAATACCGAT CTGGTCAACG
27901  AGCTGGGTGT CACCTGGGGG GCGCAGATCG GAACCACCAG CCTGGGCTAT
27951  GGCGATCTGG GGCTGCGTCC CGGCAACGGC CTGCCCGTGG ACGGCGCGGC
28001  GGCCGACCTG GCGCCCGGAA CCTTGGGGAT CAGTGTCAGT ACCCGGCTGG
28051  CGGCGCGCTT GCGTGCGTTG GAGTCGGACG GGCAGGCCAA TATCCTGTCT
28101  CAGCCGTCCA TCCTGACCGC CGACAACCTC GGCGCCATGA TAGACCTGTC
28151  GGATACCTTC TACATTCGCA CCCTGGGCGA GCGCGTAGCG ACAGTCACGC
28201  CTGTCACGGT GGGTACGTCG TTGCGTGTGA CGCCGCGCTA TATCGCCGCC
28251  AAGGGAGGAC GCCAGGTGGA ATTGGCGATC GATATCGAGG ACGGACGGGT
28301  CTTGCAGGAG TATCCCATCG ATGGTCTGCC CCGGGTTCGG AAAAGCAGCA
28351  TCAGCACGCT GGCGGTGGTG GGGGACGAGC AGACGCTGCT GATCGGCGGC
28401  TACAACAATC GCCGTGACGA AGAGCAGGTC GAGAAAGTGC CGCTGCTGGG
28451  AGATATCCCC GGCCTGGGGT TCTTGTTCTC GAGCAAGTCC CGGGCGGTAC
28501  AGCGCCGCGA GCGGCTGTTC CTGATCCGGC CGCGTGTCGT GGCTATCGAG
28551  GGCAAGCCGG TCTTCAGCCC CGTTGCGGGC ACGTCGCAGG TGTTCATGAG
28601  CACGGGTTGG GGCGGGCATG GCAGCAGCCT GAGCATTGCA CCCGGCGAGG
28651  GCGGGCATAC ACAAGTGCGT CATGATGCCC GGGCGGGCAG GCCGGTCCGG
28701  CTGGTGCCGG ATTCATTGCA TGTGGAGTAT GGCGAGGCGG GGGAGGCGTC
28751  GCCCTGAGCG TGGCCCGGGC AGGGGGGCTA CGGCACGCTG TCGTAGCCTC
28801  GTTCGCGCAG CGCTTCGCGC AAGGCACAGC GGGCGCGGGA AAGTCGGCTG
28851  CGAATGGTTC CTACAGGAAC CGAAAGCAGT GCGGCAGCCT CTTCATAGGA
28901  GAGTTCTTCC ACGCCGACCA TGAGGATCAC GTCGCGCATG CTTTCGGGGA
```

Figure 5 (continued)

28951 GCTGCTCCAG CGCTTCGCGT AGCAAGCGCA TGCGTTGACG CTGCTCGGTC

29001 ACGGCTTCGG GGTTGGGCGC ACTGCATGGC ATGACACCCA GGCTGGCGTC

29051 GGAGGTGAAT TCATAACGGC GCTCTGGCGC ACGCGACAAG TGATTGCGGA

29101 CCAGATTGAG CGCGATGCCG TACAGCCAGG TGGAAAGCTG GGAGTCGCCA

29151 CGGAACGATT GATACGCGCG CGCCGCCTCG GCGAAAGCCT GCTGCGCAAG

29201 GTCCTCGACG TCGCTGCTGT GGCCGATGTG CTTGGCGATG AATCTCTGTA

29251 ATCGGCCGGC ATGTTCGGCT ACCAGGCTGC TCAGCAACTG CTGGTCCGAT

29301 CGCTCGGTGG CGATATCGGG ATATGCGTTG TCCGGTTTTT CGAGAGAAGC

29351 GGGAAGACCG GATGACGGGG GAGAAACCAT GCAAAGCGAT ACCAAGTGAA

29401 AGGGTGATAA TTCACGTCAC CAAGATACTG ACTGCCGGTT TTATCCGGCA

29451 GTTGTTAACT TCCGAAACTA ATGTCGGATC GCGGTCGCTA CCGGAGCATT

29501 CAGATACAAC GCGCTGAACG TCTTCCGTAA AACTTACGAC GCGACGTATG

29551 GGGACTACAC CAGGTGGTGC GCAAGGACCT CGCGCAACCA TTCTTGCGCC

29601 GTATGGGCGG ATACCGGCGC CTTACCTGAT TGTGCGGCGG ATGCGAAGCG

29651 CTGCTGCATG CTTTCGTCGA TCGTGGCGCG CAACGTTCCG GATATCTGAT

29701 CGGGAGAGAC ACCGGTGTCG TGACAGACCT GTCGGAAGCC GGCGGCGTCG

29751 GCGCGGGCCG ACCAGGACAG GAAGGTAAGG AAATCCACGC CCTGCAGGGC

29801 GTGGCGTGAG GTTTCCGCCA TGTCGACCGC TCGCGTGACG ACGCGCGCCG

29851 ACAGCGATGT CGTCGGGACC GCGTGCAGAT CGAGCTCCTT GGCGACGGCG

29901 CGGGCGATTC CCGTGCCGTA GGCCAGCGCC AGCGCGCTGG AAAACATGAC

29951 AAGCGTGTCC TCGTCGGTGG CGACCCAGGA TACGTTGCGC CCTGACGGCG

30001 TCGTGCCGGA TGCGATGACC TGGAACTGCT CGCCTGCTTT GGTGACATAT

30051 AGGGTCTCTC CTTGGCTGGC GGCGCGCGCA AAGACATCAA GCTCCAAAGC

30101 AGGTAAAGCG GGTGGGATCT GGAAGTTCAT GCGGTGGGCC GTTGTCTCGA

30151 ATCCTTGGGT GTATCGCGTT CTACGGAGCG GGAACATGGA ATATGCACTG

30201 TGTGACGCTT TCGGCTCTTC GGTTCCATGC CGGCATGACA AACCCGACGG

30251 CATGCAGGCC TCGCTCCCGT CCGGCCACCG AGGCTCGCCT TGAACGCTGC

30301 GCAAAGTGTC GGCCATACGC TGGGCCTCGA GCTGGAAATG GTGGTGGCCT

30351 GTCGTGCGAC CGGGGCCAGC CATCCGGTGG CGCGCCATTT CGAAGCGCTG

30401 CGCAACTTGC GCCGTCAACG CGGAGAGTCC GTGCAGGAGT ACCGTCTGGA

30451 CGCGCGCCTG TGCGGGGGGT GGGCGGTCCC CATGGCCTGA GCGGGCTTGA

Figure 5 (continued)

```
30501  TAACGGCTAT AACCTGCTCG AAACCGCATT CGCGCCTGTG AACGGCGGCG
30551  CGGGAGGGCT GCGCCGACTG GCCGAAGCGG TGCGCCGCGA GCTGCGTGAC
30601  ACGCAGCTCG CCCTGGCCGC GGAGGGGGCG ATGCTGATCA ACGCGGCCGA
30651  GCACCCCGCC GCCAGCCTGG ATGCCGACTG GTACCGGCGA GTGCGGGTGC
30701  CGCGGCCCAT CTACGAGGAA CTCGTCGGCC AGCGAGGCTG GCTGCACCGG
30751  ATCGGGATAG ACGCCAAGGC ACAGAACAGC CCCTGCACGT CGGTTCCCGT
30801  GGCCATCGCC GCGCGCTGCC TGAACGTCGT GCTGGCGCTG GCTCCCGCGC
30851  AGATCGCCAT GTTCGCCAAC AGCCCGCTGG AGGCAGGGCG GGTGACCGGT
30901  CTCAAGGAAA ACCGCCTGAC CCTGTGGCCG CGCATGTTCC GAGGCGCGCG
30951  CTACCTGGGC GACGACCTGC TGCATCGCCT GCCTGCAAGG CCGTTTCGCG
31001  ATCTCGGCGA TTATTTCCGC TGGATGTTCG GCGGATTGAC CGCCAGCCGG
31051  GCGCTACCGC CGGGCGACGC TTGCGACTAC AAGAACGCCG ATGTGGCCTG
31101  CCTGGTGGGA GCCCCTTCGC TGGCAGAGTT CCTGTATGCG GGCGCGTGGT
31151  CCGCGCGAAA CCTGAATGAT GGCGGTTCCG TGCGTCTGGC CGCGCGCAGC
31201  GAACATTTCG TCTATTCGCA GTTCGCGCAG TTCCTGGACG CGCGTTGGCG
31251  CTACAGGATG CCGATTGTCC CCGCCTTGCC GGCGCTGTTG CGAGCCTGGG
31301  ACAGGCAGGG CGGCCTGGAA GCGCTGTTCG AGCAGGCCGG CGCGCAAGGC
31351  TACATCGAGG GGCGCGCGCC GGGCGCGGTA TTTGCCGATG CCGACTTGCT
31401  GAGCTCAGCC GGCGATGCAG TCGCGGCCAG TGCGCCGATG GCGGCGTCGG
31451  CGCTGCAATT GGGGCTGTTG CGCAATCTGC ACGACGCCGA GGCCCTGGTG
31501  AGGCGATGGG GCTGGCTGCG CTTGCGTGCG TTGCGCGATC GGGCCATCGC
31551  TTTGGCGTTG GACGATGCGC AGGTGCGCTG CCTTTGCCAA CAGGTCGTGG
31601  CGGTAGCCGA AGGCGGGCTG GCCGGCGACG AGCAGCAATG GCTCGATTAT
31651  GTGCGTTACG TGGTGGAAAC CGGCGAGACC GCCGCGGACC GCATGCTGCG
31701  CTTGTGGCGC CAGGCGCGCG GCACGCCTGA GATGCGCCGC GCACAGGCGT
31751  GCCGGCAGCG CGCGGTGCTG TCCTAGGATC CGGCACATTC CTTGCCAGGT
31801  CTCGTTGGCC AGGCCTGTCA GTGCTCCGCT TCGTACACCT CGTCGGCCGC
31851  CAGCAGAATA TCCACCGGTG GGTCGAAGCC GATGCGCCGC GCCGTTTCCA
31901  GATTGATGGC TATCTTGGCG GGGGCATTCC AGACCTGGCT GATGCTGCGC
31951  GGCTTTTCCC CATTGAAAAT GCGGGCAATG GTCTGGGCGT GGAACATGCC
```

Figure 5 (continued)

```
32001  TACGCTGGAG TAGTCCGCCT TGGCCAGGCT CATCAACAGG CCGGCCTTGA
32051  CCTCGTCGGA GCCTTGCATC GAGAAACTCG GCACGCGGGC GGCGCGCAGC
32101  AGCGCGGCGA GCTGCTTGAC GGACGTCGAG GTGATGCCCC GGTGCTCGGT
32151  GACGTAAAAG GCGTCGACTT CGCTCGACAG CTTCTGGTAG CAAGCCAGAA
32201  CGTTCTGGGT TGCCGTGGCG ATGGGGATGC CGGTCGCGCG TGCGTCGCAA
32251  CGCTTGACGG AGAAATCCAA TGCCGGCATT AGTGCGGCGA CCTTATCGAT
32301  GGCTGCGTAG GTGCGACCTG CTTCGGTGTC TTCGTAGACC AGTCCAAGCG
32351  TCTTGAACGG CACGATGTCA TGGAGCAGCT GGATCTGCCG CTGGTAGTGG
32401  TCGGGCTGTA CCCGGGCATG CAGGTTGTCC TGGCCGCTGT CGGCCGCACT
32451  GGGTATGATC CGGGCGCTTA TCGGGTCGGT CGACGAGACG ACCACGGTGG
32501  GTACCGGCGT GCCCAGTTCG ACCATGTCCT GTCCAGCCCA GGTACCCATG
32551  GCGATGATCA GGTCGATGTC CTTGGCGCCA TGCAGGCGTG CCGCAACGGC
32601  TTCGCGCACG GCAGGCCGCA AGGCGGTGTC GAAGTTGCCG GGCTGCCACC
32651  ACGCATCGGG CACGAACTCG ATGTAGTTGC TGCGGGCATG CGTGGCCAGG
32701  TAAAGCCAGG CCTTTCGCAT ATCGGTTATC TCGGGCATGT CGTCGATACG
32751  CAGCCATCCG AGTTGTTGCA ATGCGCGCGC GATCGCGTAG AGCGTGCGCG
32801  GATACTCCTC GTACTCGCCG CTACCCACAT AACCGATGCG CCATTTCCGG
32851  CCGGATGTAT GGGAGGGAGG CGGAAGGCGG GGCGAGGTAA GGGCGACGCC
32901  GGAGCTCAGG GCCGCGACAG GAGGAGGGCT GGATGCCGCC GCGATGGGCC
32951  ACGCGAGGCC AAGAAGCAGG GCGAGCGGGG CGAGTATTCC GGGGCGAAGG
33001  GTCATGGGCG ATGAATGGCG ATGATGGTGA GATCGTCGGA TTGTTCGGAA
33051  TCGGCGGCGA ATTCGCGTAG GTCGTGCAGA ATGTGCTCGA TGAGTTCGGC
33101  CGCTGCGTGC GGCGCGCCCT GCATCAGGGC GACCAGCCGC GGCAGACCAT
33151  ACTGGGCGCA GCCGCCGTGG ATGGCTTCGG TGACGCCGTC GGTAAACGCG
33201  ACCAGCGAGG TGCCGTTCGG CAAGGTGGTG CTCAGGGTGG AATACGCCTC
33251  GTTGTCCAGC ACGCCGCAGG CCGCGCCGCT GCTTCCTTGA AGCAGGCGGA
33301  CCTCGCCACG TTCGTCGATG AGCAGCGGCG GCGGGTGGCC GGCGTTGACC
33351  CAGGCCAGGG CGCCTGTTTC CGGGGTGAAG ACGCCTATCA GCAAGGTGAC
33401  AAACATCAGC TTGGGGTTGT TCTCGGCCAG ACGGTGGTTC ACCTTGGTGG
33451  CGATGGCGCC CGGGTCGTGC TCTTCTTCCG CCACGCTGCG TATCAAGGTC
33501  CTGACGATGG CCATGAACAG GGCCGCGGGC ACGCCTTTTC CGGATACGTC
```

Figure 5 (continued)

```
33551  GCCGATGGCA AAGCACAGAC GCCCGTCTGC CAGCACGAAG TAGTCGTAGA
33601  AATCCCCACC GACCTCCCGG GCCGGGTACA TGACGGCACG CAACTGGCTG
33651  CCGCGCGTGG CCGCATCGGG CAACGGCTGG GGAAGCAGGC CAAGTTGGAT
33701  GGAGCGGGCG ATGCTCAATT CGCTTTCGAG GCGTTCGCGG TTCGATATCT
33751  GCGCCATCAG GGCCCGCACA TTGTGGTGCA GCTGTTCGTT CATGAACAGG
33801  AACGATTCGG CGAGCTGTCC GACTTCGTCG CGCCGCCGGC GCGGCAGGCA
33851  TGCCACCGAC GGCGGAACCC GGATCGGCTC GGTGAGGTCC TGGGTGGGAA
33901  GCTGGCGAGC GTAGTTGCTC AGTTGCGCCA ACGGCCGGGC GATGCGCACC
33951  GCCACCACCC ATGCCAGCAT CAGCCCGGCC AGCAAGGTGG CGGCGAAGAT
34001  CAGTGCCTGC CGGCGCACCA GATTCTGTGC CGGGTCGGTC AGGTCCGGCT
34051  CGGGAACGAC ACCGATGATG GTCCAATGCA GCGGCTTGTA TCGCAGGGCG
34101  TCGATCTGCC AGGCGCTTTC GCCGTTGGTA AAGCGCAACG TCAGGCCGCG
34151  GGTAGACGAG ATTTCGGCAA GCATCGAATG CAATACCCGT CCCGATTCGA
34201  CGTCTGTCGA GTCCAGCAGC CGGGCGGCCG ATGGGGGTGG CGGCACGATC
34251  ACCGTGCCAT CGTCCGCAAC CACGAACACG AAACCATGGC GGCTGAGCCG
34301  CAGCTCCGAC AGGTTCCGGT CTATCGCGGC AATCATGTTG GCTTTCTGGG
34351  CGGCAACCTT GTCGATGATG GCTTGTGAGC TATCGGAGAT GGCGAGAACC
34401  CACTTCCACG CGGGGAAGTA CACGAAATAG GCGTGTCGCA TCTGGGCGGA
34451  CTCGTCCAGG GGAGACGGGT AGATGGCGAA GCCGCGACCG TCGTTGCGGC
34501  TTTCCTCGTA CATGGCGGCG GCGAGCGGCC GGCCCTTGAA GTCGCGGATC
34551  CCGGAGAGGT CCCGGTCGAT CATCCGGGGG TTGGTGCTGG CCAGCACGGT
34601  GCCTTCCGCG TCATAGGCGA AGGCGACGCG GCGCGGTCCC AGGTCGAGAT
34651  GGTTCAGCCA GACACGCGCC ATGCCCTTGG CGGCGCCGGT AGTGACGTGT
34701  CCGCGCTCGG CCTGCGCGGC ATAGGCGTTC AGCACCGATG TCACGACCGC
34751  GCTGAGTTGG ATCAGTTGCC TGCGGCTTTC CCGGATGGTG CGGATCTTGT
34801  CGTCGAGCAG CGTGGACCAG CGCGTGTCGG TGTCACGAAC TACCAAGTCC
34851  ATGATGTTAC TGACGGCATG CAGTTCGTTC TTGATGATGT TGTTCGTGAC
34901  ATCGCGCTGG GTCACGAGCA TCACGACGAT GCCTACGAGT AGCAGCGTGG
34951  ATGCAATGAG CAGGAGGAAC TTCCCACGCA ATGAAAGCGG CAACTCTAGC
35001  CGGGGAGTAC GGCGCATGAA CATGAA
```

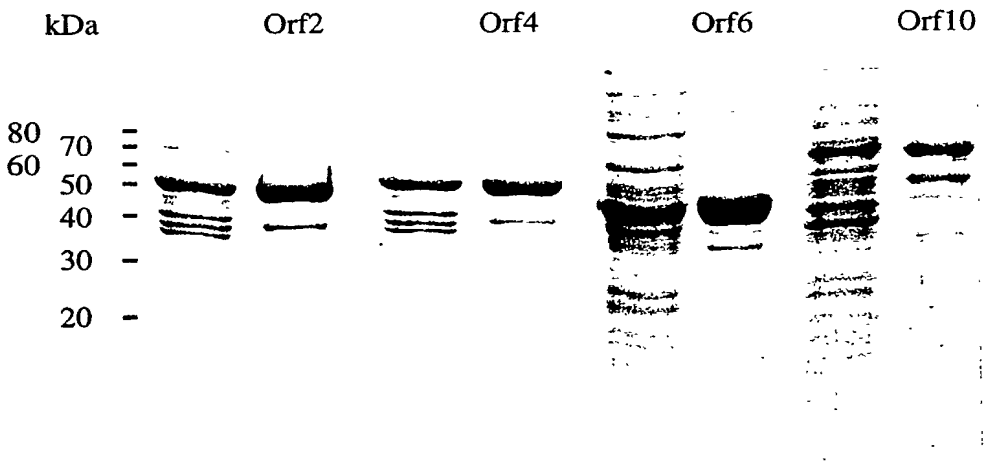

FIGURE 6

Docket No.: B45168

PCT/EP99/10297

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

VACCINE

the specification of which (check one)

[ ] is attached hereto.

[X] was filed on 21 December 1999 as Serial No. PCT/EP99/10297
and was amended on (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or Inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)

Number	Country	Filing Date	Priority Claimed
9828217.1	Great Britain	21 December 1998	Yes

I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below.

Application Number	Filing Date

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

Serial No.	Filing Date	Status

I hereby appoint the practitioners associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

Customer Number 20462.

Address all correspondence and telephone calls to Zoltan Kerekes , SmithKline Beecham Corporation, Corporate Intellectual Proprety-U.S., UW2220, P.O. Box 1539, King of Prussia, Pennsylvania 19406-0939, whose telephone number is 610-270-5024.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00

Full Name of Inventor: Alex BOLLEN

Inventor's Signature: _____ Date: June 14th 2001

Residence: Brussels, Belgium BEX

Citizenship: Belgian

Post Office Address: SmithKline Beecham Corporation
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, Pennsylvania 19406-0939

2-00

Full Name of Inventor: Alain FAUCONNIER

Inventor's Signature: _____ Date: June 18th 2001

Residence: Brussels, Belgium BEX

Citizenship: Belgian

Post Office Address: SmithKline Beecham Corporation
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, Pennsylvania 19406-0939

3-00

Full Name of Inventor: Edmond GODFROID

Inventor's Signature: _____ Date: 14-05-01

Residence: Brussels, Belgium BEX

Citizenship: Belgian

Post Office Address: SmithKline Beecham Corporation
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, Pennsylvania 19406-0939